

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 17:12:26 ; Search time 5147 Seconds
(without alignments)
10252.133 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcggcagcgccgca.....aaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4708233 seqs, 24227607955 residues

Word size : 0

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sta.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965	88.6	1089	6 AX100704	AX100704 Sequence
2	648	59.5	1083	6 TSP271917	AJ271917 Triticum
3	168	15.4	168	6 AX100712	AX100712 Sequence
4	168	15.4	168	6 AX100714	AX100714 Sequence
5	153	14.0	153	6 AX100710	AX100710 Sequence
6	150	13.8	150	6 AX100716	AX100716 Sequence
7	114	10.5	156	6 AX100708	AX100708 Sequence
8	60	5.5	1132	8 SLU39059	U39059 Lycopersico
9	59	5.4	1630	9 BC028028	BC028028 Homo sapi
10	57	5.2	246093	2 CR628367	CR628367 Mus muscu
11	56	5.1	444	9 BC002750	BC002750 Homo sapi
12	56	5.1	1980	9 BC015791	BC015791 Homo sapi
13	56	5.1	1985	3 AX118692	AX118692 Drosophil
14	56	5.1	2084	9 AB072761	AB072761 Macaca fa
15	56	5.1	3430	9 BC065220	BC065220 Homo sapi
16	56	5.1	3441	10 BC038365	BC038365 Mus muscu
17	56	5.1	3606	9 BC030162	BC030162 Homo sapi
18	56	5.1	4430	5 BC059794	BC059794 Danio rer
19	56	5.1	4669	5 BC048022	BC048022 Xenopus 1

20	5.1	114458	10	BX293986	Mouse DNA
21	5.1	161155	2	AC128086	Rattus no
22	5.1	179678	2	AC123687	Mus muscu
23	5.1	208225	2	AC110537	Mus muscu
24	5.1	237538	2	AC108561	Rattus no
25	5.1	244978	2	AC128734	Rattus no
26	5.1	376	6	CQ526562	Sequence
27	5.1	491	6	AR487921	Sequence
28	5.1	764	5	AF354917	Trimeresu
29	5.1	798	9	AK026442	Homo sapi
30	5.1	833	9	AK024546	Homo sapi
31	5.1	840	5	BC083489	Danio rer
32	5.1	850	9	BC036763	Homo sapi
33	5.1	888	9	BC046163	Homo sapi
34	5.1	1489	10	BC028982	Mus muscu
35	5.1	1665	10	BC061850	Rattus no
36	5.1	1723	5	BC068376	Danio rer
37	5.1	1789	9	AK074236	Homo sapi
38	5.1	2021	9	BC008075	Homo sapi
39	5.1	2528	3	AY051822	Drosophil
40	5.1	2673	10	BC0211308	Mus muscu
41	5.1	2727	5	BC067323	Xenopus t
42	5.1	2859	10	BC033303	Mus muscu
43	5.1	2973	9	BC042518	Homo sapi
44	5.1	3422	9	BSM805920	Homo sapi
45	5.1	3819	6	CQ768015	Sequence

ALIGNMENTS

RESULT 1	AX100704	Sequence 1 from Patent WO0121644.	1089 bp	DNA	linear	PAT 10-APR-2001
LOCUS	AX100704	AX100704.1	GI:13619652			
DEFINITION	AX100704	Triticum monococcum				
ACCESSION	AX100704	Triticum monococcum				
VERSION	AX100704.1	Triticum monococcum				
KEYWORDS		Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae; Triticeae; Triticum.				
SOURCE		Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.				
ORGANISM		Wheat gp proteins and uses thereof				
REFERENCE		Patent: WO 0121644-A 1 29-MAR-2001;				
AUTHORS		CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)				
TITLE		Location/Qualifiers				
JOURNAL		1. .1089				
FEATURES		/organism="Triticum monococcum"				
source		/mol_type="unassigned DNA"				
		/db_xref="taxon:4568"				
		20. .805				
		/codon_start=1				
		/protein_id="CAC36466.1"				
		/db_xref="GI:13619653"				
		/translation="MAPPRGAAATAALDLTGVIHLEASSVPPLPERGGNAVORRG				
		AVDPDKRKKKAAAPRITGRLREYSKIVCEKVEAKGRITTYNEADEIYSELKMAH				
		IQGDFEKIRRRVDAFNVLIAKEKEIKRWGLSNRYEKIKKLEEVKELV				
		NKIRNKALLOIEKQFDLQNIKLNOTLESSAENVNGIRLPFLVLKTSRKARVEIE				
		ISDDSKFAHFENFNGAFTLHDDLISILEGVRRNSIGRATLH"				

ORIGIN

Query Match	88.6%;	Score 965;	DB 6;	Length 1089;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches	965;	Conservative 0;	Mismatches 0;	Indels 0;
Gaps	0;			
QY	125	GGCGGTATGCGGTCCAAAGGAAGGGGCTGTGTACCCGGATAAAGATAGGAAGAGGAG	184	
Db	125	GGCGGTATGCGGTCCAAAGGAAGGGGCTGTGTACCCGGATAAAGATAGGAAGAGGAG	184	

```
QY 185 AAGGCTGCGGACCGAGGATCACCGTTCGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244
Db 185 AAGGCTGCGGACCGAGGATCACCGTTCGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244
QY 245 GAGAAAGTTGAAGCCAAAGGAGAAACAACATACAATGAGGTTCGACAGAAATTTATTCA 304
Db 245 GAGAAAGTTGAAGCCAAAGGAGAAACAACATACAATGAGGTTCGACAGAAATTTATTCA 304
QY 305 GAGCTGAAGTCCATGCGACATATTTGTCAGAGGTTTGTAGAGAAATATTAGCGCGAGA 364
Db 305 GAGCTGAAGTCCATGCGACATATTTGTCAGAGGTTTGTAGAGAAATATTAGCGCGAGA 364
QY 365 GTGTATGATGCTTTCACAGTCTCTCAATGACATTCGTTGTTATTCGAAAGAAAGAGAG 424
Db 365 GTGTATGATGCTTTCACAGTCTCTCAATGACATTCGTTGTTATTCGAAAGAAAGAGAG 424
QY 425 ATACGCTGGATGGCGCTTTCAAATTAACAGATATGAAAAATAAAGAGCTTCGAGGAAGTT 484
Db 425 ATACGCTGGATGGCGCTTTCAAATTAACAGATATGAAAAATAAAGAGCTTCGAGGAAGTT 484
QY 485 CGTAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGAAAAA 544
Db 485 CGTAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGAAAAA 544
QY 545 CAGTTTGTATGATCTCCAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604
Db 545 CAGTTTGTATGATCTCCAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAG 604
QY 605 AATGTTAATGGCATCGCCCTTCATTCGTTTGGTCAAGACATCTAGGAAAGCAAGGGTG 664
Db 605 AATGTTAATGGCATCGCCCTTCATTCGTTTGGTCAAGACATCTAGGAAAGCAAGGGTG 664
QY 665 GAAATGAGATTCAGATGATCGAAGTTTGCCCATTTGAGTTCAATGGTGCACCATTC 724
Db 665 GAAATGAGATTCAGATGATCGAAGTTTGCCCATTTGAGTTCAATGGTGCACCATTC 724
QY 725 ACATTGCATGATCTCTCAATCTTGGGGGTAGGCGTAACAGCATAGGAAGAGCT 784
Db 725 ACATTGCATGATCTCTCAATCTTGGGGGTAGGCGTAACAGCATAGGAAGAGCT 784
QY 785 GGCGCGGCCACCCCTTCACTAGAGACTCAAGAATATTACAAATGAATTAAGAGTTTAGAA 844
Db 785 GGCGCGGCCACCCCTTCACTAGAGACTCAAGAATATTACAAATGAATTAAGAGTTTAGAA 844
QY 845 CTGGCAGCGCGATCTTTTGCACAGCTATGATAGCTATATATCTCATGAAACTTG 904
Db 845 CTGGCAGCGCGATCTTTTGCACAGCTATGATAGCTATATATCTCATGAAACTTG 904
QY 905 ACCTAGTTTATAGCAGCTCTCTCAGGCTTGAGAAGATTTTACCTGCAAAATTTGCTC 964
Db 905 ACCTAGTTTATAGCAGCTCTCTCAGGCTTGAGAAGATTTTACCTGCAAAATTTGCTC 964
QY 965 CTTTTTGTGCTAGCAGGTTATTAGGTCTCAGATAGATGATTCATATATGCTGCTATG 1024
Db 965 CTTTTTGTGCTAGCAGGTTATTAGGTCTCAGATAGATGATTCATATATGCTGCTATG 1024
QY 1025 ABAACATTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084
Db 1025 ABAACATTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084
QY 1085 AAAAA 1089
Db 1085 AAAAA 1089
```

```
RESULT 2
TSP271917
LOCUS Triticum sp. mRNA for DP protein (dp gene). linear PLN 16-DEC-2000
DEFINITION Triticum sp. mRNA for DP protein (dp gene).
ACCESSION AJ271917
VERSION AJ271917.1 GI:11877790
KEYWORDS DP gene; E2F dimerization partner.
SOURCE Triticum sp.
```

```
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL
FEATURES
source
gene
5'UTR
CDS
3'UTR
ORIGIN
Query Match 59.5%; Score 648; DB 8; Length 1083;
Best Local Similarity 99.6%; Pred. No. 0;
Matches 798; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

```
QY 125 GGCGGTAAATGCGGTCCAAAGGAGGCGGTGTTGACCCGGATAAAGATAGGAAGAGGAG 184
Db 119 GGCGGTAAATGCGGTCCAAAGGAGGCGGTGTTGACCCGGATAAAGATAGGAAGAGGAG 178
QY 185 AAGGCTGCGGACCGAGGATCACCGTTCGGGGCTCCGGAGTACAGCAAAATAGTTTGT 244
Db 179 AAGGCTGCGGACCGAGGATCACCGTTCGGGGCTCCGGAGTACAGCAAAATAGTTTGT 238
QY 245 GAGAAAGTTGAAGCCAAAGGAGAAACAACATACAATGAGGTTCGACAGAAATTTATTCA 304
Db 239 GAGAAAGTTGAAGCCAAAGGAGAAACAACATACAATGAGGTTCGACAGAAATTTATTCA 298
QY 305 GAGCTGAAGTCCATGCGACATATTTGTCAGAGGTTTGTAGAGAAATATTAGCGCGAGA 364
Db 299 GAGCTGAAGTCCATGCGACATATTTGTCAGAGGTTTGTAGAGAAATATTAGCGCGAGA 358
QY 365 GTGTATGATGCTTTCACAGTCTCTCAATGACATTCGTTGTTATTCGAAAGAAAGAGAG 424
Db 359 GTGTATGATGCTTTCACAGTCTCTCAATGACATTCGTTGTTATTCGAAAGAAAGAGAG 418
QY 425 ATACGCTGGATGGCGCTTTCAAATTAACAGATATGAAAAATAAAGAGCTTCGAGGAAGTT 484
Db 419 ATACGCTGGATGGCGCTTTCAAATTAACAGATATGAAAAATAAAGAGCTTCGAGGAAGTT 478
QY 485 CGTAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGAAAAA 544
|||||
```

Triticum sp.
Eukaryota, Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.

1 Ramirez-Parra, E. and Gutierrez, C.
Characterization of wheat DP, a heterodimerization partner of the
Plant E2F transcription factor which stimulates E2F-DNA binding
FEBS Lett. 486 (1), 73-78 (2000)

20562804
11108846
2 (bases 1 to 1083)
Gutierrez, C.
Direct Submission

Submitted (03-FEB-2000) Gutierrez C., C.S.I.C.-U.A.M., Centro de
Biologia Molecular, Cantoblanco, Madrid 28049, SPAIN
Location/Qualifiers

1. .1083
/organism="Triticum sp."
/mol_type="mRNA"
/db_xref="taxon:4569"
1. .1083
/gene="dp"
1. .13
/gene="dp"
14. .799
/gene="dp"
/function="cell cycle E2F dimerization partner"

/codon_start=1
/product="DP protein"
/protein_id="Cac19034.1"
/db_xref="GI:11877791"
/db_xref="GOA:Q9FET1"
/db_xref="UniProt/TREMBL:Q9FET1"
/translation="MAPRGGAAATAALDLTGWHILEASSVPPPERGNAVQRKQ
AVDPKDKKERRVDAFNLIALRVIAKEKEIRWGLSNRYEKIKLEEVKELV
NKIRKKALLOIEKQFDLONI KLRNOTLESSAENVNGIRLPFVLVTKSRKARVEI
ISDDSKPAHFENFNGAPFTLHDLDSILEGVRGNSIGKAGRAITLH"
801. .1033
/gene="dp"

```
Db 479 CGTAAGAAGTCTGCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 538
QY 545 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAAGCTCAGCAGAG 604
Db 539 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAAGCTCAGCAGAG 598
QY 605 AATGTTAATGGGATCGGCTTCATTCGTTATGGTCAAGACATCTAGGAAGCAAGGGTG 664
Db 599 AATGTTAATGGGATCGGCTTCATTCGTTATGGTCAAGACATCTAGGAAGCAAGGGTG 658
QY 665 GAAATTTGAGATTTTCAGATGACTCGAAGTTTGCCCATTTTCAGTTCAATGGTGCACCATTC 724
Db 659 GAAATTTGAGATTTTCAGATGACTCGAAGTTTGCCCATTTTCAGTTCAATGGTGCACCATTC 718
QY 725 ACATTCATGATGATCTCAATTCCTTTCAGGCGGTGAAGCGTAAACAGCATAGGAAGAGCT 784
Db 719 ACATTCATGATGATCTCAATTCCTTTCAGGCGGTGAAGCGTAAACAGCATAGGAAGAGCT 778
QY 785 GCGCGCGCCACCTTCCTAGTACACTCAAGATATTTACAAATGAATTAAGTTGTAGAA 844
Db 779 GCGCGCGCCACCTTCCTAGTACACTCAAGATATTTACAAATGAATTAAGTTGTAGAA 838
QY 845 CTGGCACAGCGGATTCCTTTGACACAGCTATGTATAGCTATATATCTCATGAAGAACTTG 904
Db 839 CTGGCACAGCGGATTCCTTTGACACAGCTATGTATAGCTATATATCTCATGAAGAACTTG 898
QY 905 ACCTAGTTTATAGGACAGTCT 925
Db 899 ACCTAGTTTATAGGACAGTCT 919

RESULT 3
AX100712
LOCUS 168 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 9 from Patent WO0121644.
ACCESSION AX100712
VERSION AX100712.1 GI:13619660
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 9 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source 1..168
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
CDS
1..>168
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36470.1"
/db_xref="GI:13619661"
/translation="MAHICGGFDEKNIRRVYDAFNVLIALRVIAKEKKEIRWMGLSN
YRYEKIKLLEEV"

ORIGIN
Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 376
Db 1 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 60
QY 377 TTCAACGTTCTCATTCGACTTCCTGTTATTGCAAGAAAGAAAGAGATACGGTGGATG 436
Db 61 TTCAACGTTCTCATTCGACTTCCTGTTATTGCAAGAAAGAAAGAGATACGGTGGATG 120

Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 376
Db 1 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 60
QY 377 TTCAACGTTCTCATTCGACTTCCTGTTATTGCAAGAAAGAAAGAGATACGGTGGATG 436
Db 61 TTCAACGTTCTCATTCGACTTCCTGTTATTGCAAGAAAGAAAGAGATACGGTGGATG 120
```

```
QY 437 GGCCTTTCAAATTACAGATATGAAAAATAAAGAGCTTTGAGGAAGTT 484
Db 121 GGCCTTTCAAATTACAGATATGAAAAATAAAGAGCTTTGAGGAAGTT 168

RESULT 4
AX100714
LOCUS 168 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 11 from Patent WO0121644.
ACCESSION AX100714
VERSION AX100714.1 GI:13619662
KEYWORDS Triticum monococcum
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 11 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source 1..168
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
CDS
1..>168
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36471.1"
/db_xref="GI:13619663"
/translation="RKELVNKIRNKKALLQEIEKQFDLQIKLRNQTLSSAENVNG
IRLPFLVKTSR"

ORIGIN
Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 CGTAAGAAGTCTGTCACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 544
Db 1 CGTAAGAAGTCTGTCACAGATTAGGAACAAGAGGCACTCTCCAGGAATCGAAAA 60
QY 545 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAAGCTCAGCAGAG 604
Db 61 CAGTTTGATGATCTCAAAACATCAAGTTACGTAACCAACACTGGAAAGCTCAGCAGAG 120
QY 605 AATGTTAATGGCATCGGCTTCATTCGTTATGGTCAAGACATCTAGG 652
Db 121 AATGTTAATGGCATCGGCTTCATTCGTTATGGTCAAGACATCTAGG 168

RESULT 5
AX100710
LOCUS 153 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 7 from Patent WO0121644.
ACCESSION AX100710
VERSION AX100710.1 GI:13619658
KEYWORDS Triticum monococcum
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 7 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source 1..153
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
CDS
1..>153
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36470.1"
/db_xref="GI:13619661"
/translation="MAHICGGFDEKNIRRVYDAFNVLIALRVIAKEKKEIRWMGLSN
YRYEKIKLLEEV"

ORIGIN
Query Match 15.4%; Score 168; DB 6; Length 168;
Best Local Similarity 100.0%; Pred. No. 5.2e-79;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 376
Db 1 ATGGCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCT 60
QY 377 TTCAACGTTCTCATTCGACTTCCTGTTATTGCAAGAAAGAAAGAGATACGGTGGATG 436
Db 61 TTCAACGTTCTCATTCGACTTCCTGTTATTGCAAGAAAGAAAGAGATACGGTGGATG 120
```

```

/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<!.>153
/codon_start=1
/notes="unnamed protein product"
/protein_id="CAC36469.1"
/db_xref="GI:13619659"
/translation="DKDRKKEKAAAPRIITGWGLREYSKIVCEKVEAKGRTTYNEVADE
IYSELKS"

ORIGIN
Query Match      14.0%; Score 153; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 6.4e-71;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 GATAAAGATAGGAAGAGGAGGCTGGCGCACCGAGGATCACCGTTGGGGGCTCCGC 223
Db 1 GATAAAGATAGGAAGAGGAGGCTGGCGCACCGAGGATCACCGTTGGGGGCTCCGC 60

QY 224 GAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAG 283
Db 61 GAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAG 120

QY 284 GTTGCAGACCAAAATTTATTTCAGAGCTGAAGTCC 316
Db 121 GTTGCAGACCAAAATTTATTTCAGAGCTGAAGTCC 153

RESULT 6
LOCUS AX100716 150 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 13 from Patent WO0121644.
ACCESSION AX100716
VERSION AX100716.1 GI:13619664
KEYWORDS Triticum monococcum
ORGANISM Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
Wheat dp proteins and uses thereof
Patent: WO 0121644-A 13 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source
1. .150
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<!.>150
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36472.1"
/db_xref="GI:13619665"
/translation="KARVEIEISDDSKFAHFEFNGAPPTLHDDLILGVRNSIGRA
GRATLH"

CDS
LOCUS SLU39059 1132 bp mRNA linear PLN 18-NOV-1996
DEFINITION Lycopersicon esculentum myo-inositol monophosphatase 3 mRNA,
complete cds.
ACCESSION U39059
VERSION U39059.1 GI:1098970
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Plant inositol monophosphatase is a lithium-sensitive enzyme
encoded by a multigene family
Plant Cell 7 (12), 2175-2185 (1995)
JOURNAL MEDLINE 96351935
PUBMED 8718627
REFERENCE 2 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Direct Submission
AUTHORS Submitted (20-OCT-1995) Glenda E. Gillaspay, Plant Biology,
University of California, 111 Koshland Hall, Berkeley, CA 94720,
JOURNAL

QY 713 GGTGCACCATTCACATGATGATCTCAATCCTTGAGGGGTAAGCGGTACACG 772
Db 61 GGTGCACCATTCACATGATGATCTCAATCCTTGAGGGGTAAGCGGTACACG 120

QY 773 ATAGGAAGAGCTGGCGCGCCACCCCTTCAC 802

```

```

Db 121 ATAGGAAGAGCTGGCGCGCCACCCCTTCAC 150

RESULT 7
LOCUS AX100708 156 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 5 from Patent WO0121644.
ACCESSION AX100708
VERSION AX100708.1 GI:13619656
KEYWORDS Triticum monococcum
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
Wheat dp proteins and uses thereof
Patent: WO 0121644-A 5 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source
1. .156
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<!.>156
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC36468.1"
/db_xref="GI:13619657"
/translation="ARAAVAPRGAAAAATAALDTGVHILEASSVPLPERGNAV
QRKGAVDP"

CDS
LOCUS SLU39059 1132 bp mRNA linear PLN 18-NOV-1996
DEFINITION Lycopersicon esculentum myo-inositol monophosphatase 3 mRNA,
complete cds.
ACCESSION U39059
VERSION U39059.1 GI:1098970
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Plant inositol monophosphatase is a lithium-sensitive enzyme
encoded by a multigene family
Plant Cell 7 (12), 2175-2185 (1995)
JOURNAL MEDLINE 96351935
PUBMED 8718627
REFERENCE 2 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Direct Submission
AUTHORS Submitted (20-OCT-1995) Glenda E. Gillaspay, Plant Biology,
University of California, 111 Koshland Hall, Berkeley, CA 94720,
JOURNAL

QY 8 GCACGAGCGCAATGGCGCTCCCGCGCGAGCTGTCGGCGCGCTACCGCGGCACTG 67
Db 1 GCACGAGCGCAATGGCGCTCCCGCGCGAGCTGTCGGCGCGCTACCGCGGCACTG 60

QY 68 GACCTGACCGCGGTGCACATTCGAACTCCAGTGTCCCGCTCCCGAA 121
Db 61 GACCTGACCGCGGTGCACATTCGAACTCCAGTGTCCCGCTCCCGAA 114

RESULT 8
LOCUS SLU39059 1132 bp mRNA linear PLN 18-NOV-1996
DEFINITION Lycopersicon esculentum myo-inositol monophosphatase 3 mRNA,
complete cds.
ACCESSION U39059
VERSION U39059.1 GI:1098970
KEYWORDS Lycopersicon esculentum (tomato)
SOURCE Lycopersicon esculentum
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Solanum; Lycopersicon.
1 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Plant inositol monophosphatase is a lithium-sensitive enzyme
encoded by a multigene family
Plant Cell 7 (12), 2175-2185 (1995)
JOURNAL MEDLINE 96351935
PUBMED 8718627
REFERENCE 2 (bases 1 to 1132)
Gillaspay, G.E., Keddle, J.S., Oda, K. and Gruissem, W.
Direct Submission
AUTHORS Submitted (20-OCT-1995) Glenda E. Gillaspay, Plant Biology,
University of California, 111 Koshland Hall, Berkeley, CA 94720,
JOURNAL

```


humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk
 On Sep 28, 2004 this sequence version replaced gi:51966191.
 Sequence from the Mouse Genome Sequencing Consortium whole genome
 shotgun may have been used to confirm this sequence. Sequence data
 from the whole genome shotgun alone has only been used where it has
 a phred quality of at least 30.
 ----- Genome Center
 Center: Wellcome Trust Sanger Institute
 Center code: SC
 Web site: http://www.sanger.ac.uk
 Contact: humquery@sanger.ac.uk
 ----- Project Information
 Center project name: bm75B12
 ----- Summary Statistics
 Assembly program: XGAP4; version 4.5
 Chemistry: Dye-terminator; 100% of reads
 Consensus quality: 245844 bases at least Q40
 Consensus quality: 245914 bases at least Q30
 Consensus quality: 245977 bases at least Q20
 Insert size: 245993; sum-of-contigs
 Insert size: 220507; 3.1% error; agarose-fp
 Quality coverage: 10.26x in Q20 bases; sum-of-contigs Quality
 coverage: 12.01x in Q20 bases; agarose-fp

 * NOTE: This is a 'working draft' sequence. It currently
 * consists of 2 contigs. The true order of the pieces
 * is not known and their order in this sequence record is
 * arbitrary. Gaps between the contigs are represented as
 * runs of N, but the exact sizes of the gaps are unknown.
 * This record will be updated with the finished sequence.
 * as soon as it is available and the accession number will
 * be preserved.
 *
 * 1 132816: contig of 132816 bp in length
 * 132817 132916: gap of 100 bp
 * 132917 246093: contig of 113177 bp in length.
 FEATURES
 source
 1. .246093
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 /chromosome="X"
 /clone_lib="RP23-75B12"
 /clone_lib="RPCT-23"
 misc_feature
 1. .132816
 /notes="assembly fragment:04407
 fragment chain:1"
 132917. .246093
 /notes="assembly fragment:04662
 fragment chain:1"
 clone_end:SP6
 vector_side:right
 ORIGIN
 Query Match 5.2%; Score 57; DB 2; Length 246093;
 Best Local Similarity 100.0%; Pred. No. 2.9e-19;
 Matches 57; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1033 GATAGCAA 1089
 |||
 DB 180762 GATAGCAA 180818
 |||
 RESULT 11
 BC002750
 LOCUS Homo sapiens chromosome 11 open reading frame 10, mRNA (cdna clone
 DEFINITION MGC:3384 IMAGE:3633279), complete cds.
 ACCESSION BC002750
 VERSION BC002750.2 GI:38014315
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE
 AUTHORS
 JOURNAL
 TITLE
 JOURNAL
 PUBMED
 12477932
 2 (bases 1 to 444)
 Strausberg,R.
 Direct Submission
 Submitted (05-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 On Oct 28, 2003 this sequence version replaced gi:12803818.
 Contact: MGC help desk
 Email: cgapps-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mgc@nih.gov
 Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
 Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 12 Row: k Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7656933.
 Location/Qualifiers
 1. .444
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MGC:3384 IMAGE:3633279"
 /tissue_type="Uterus, endometrium adenocarcinoma"
 /clone_lib="NIH MGC 44"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1. .444
 /gene="Clorf10"
 /db_xref="LocusID:746"
 27. .266
 /gene="Clorf10"
 /codon_start=1
 gene
 CDS

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 444)
 Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
 Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
 Althaus,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
 Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
 Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
 Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
 Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
 Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
 Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
 McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
 Worley,K.C., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
 Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
 Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
 Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
 Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
 Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
 Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
 Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 12477932
 2 (bases 1 to 444)
 Strausberg,R.
 Direct Submission
 Submitted (05-FEB-2001) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 On Oct 28, 2003 this sequence version replaced gi:12803818.
 Contact: MGC help desk
 Email: cgapps-r@mail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mgc@nih.gov
 Akhter,N., Ayale,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
 Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 12 Row: k Column: 24
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7656933.
 Location/Qualifiers
 1. .444
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MGC:3384 IMAGE:3633279"
 /tissue_type="Uterus, endometrium adenocarcinoma"
 /clone_lib="NIH MGC 44"
 /lab_host="DH10B-R"
 /note="Vector: pOTB7"
 1. .444
 /gene="Clorf10"
 /db_xref="LocusID:746"
 27. .266
 /gene="Clorf10"
 /codon_start=1
 gene
 CDS

/product="chromosome 11 open reading frame 10"

/protein_id="AAH02750.1"

/db_xref="GI:12803819"

/db_xref="LocusID:746"

/translation="MELEWSRYTSPNPAVPHLTVLLAIGMPTFAWFFVYEVST
KYTRDIYKELLISLVASLPMFGVLFLLLVGIYV"

ORIGIN

Query Match 5.1%; Score 56; DB 9; Length 444;

Best Local Similarity 100.0%; Pred. No. 1.3e-18;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089

Db 377 ATAGCAA 432

RESULT 12

BC015791

LOCUS

DEFINITION Homo sapiens, clone IMAGE:4866083, mRNA.

ACCESSION BC015791

VERSION BC015791.1 GI:16041805

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1980)

Strausberg,R.

Direct Submission

Submitted (09-OCT-2001) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Genome Sequence Centre,

BC Cancer Agency, Vancouver, BC, Canada

info@bcgsc.bc.ca

Steven Jones, Jennifer Aeano, Ian Bosdet, Yaron Butterfield,

Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,

Leticia Hailao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo

Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven

Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline

Schein, Duane Smalls, Michael Smith, Lorraine Spence, Jeff Stott,

Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,

George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAL Plate: 34 Row: j Column: 17

This clone was selected for full length sequencing because it

passed the following selection criteria: Hexamer frequency ORF

analysis.

FEATURES

source

1. .1980

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4866083"

/tissue_type="Colon, adenocarcinoma"

/clone_lib="NIH MGC 15"

/lab_host="DH10B-R"

/note="Vector: pOT37"

ORIGIN

Query Match

Best Local Similarity 100.0%; Pred. No. 1.2e-18;

5.1%; Score 56; DB 9; Length 1980;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089

Db 1914 ATAGCAA 1969

RESULT 13

AY118692

LOCUS

DEFINITION Drosophila melanogaster AT17414 full insert cDNA.

ACCESSION AY118692

VERSION AY118692.1 GI:21429747

KEYWORDS

SOURCE Drosophila melanogaster (fruit fly)

ORGANISM

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 1985)

REFERENCE

AUTHORS

Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,

Champe,M., Chavez,C., Dorsett,V., Dresnek,D., Farfan,D., Friese,E.,

George,R., Gonzalez,M., Guarin,H., Kronmiller,B., Li,P., Liao,G.,

Miranda,A., Mungall,C.J., Nunoo,J., Pacleb,J., Paragas,V., Park,S.,

Patel,S., Phouanavong,S., Wan,K., Yu,C., Lewis,S.E., Rubin,G.M.

and Celniker,S.

Direct Submission

TITLE

JOURNAL

COMMENT

Sequence submitted by:

Berkeley Drosophila Genome Project

Lawrence Berkeley National Laboratory

Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to

sequence clones from Drosophila Gene Collection 1 (Rubin et al.,

Science 2000). The sequence has been subjected to integrity checks

for sequence accuracy, presence of a polyA tail and contiguity

within 100 kb in the genome. Thus we believe the sequence to

reflect accurately this particular cDNA clone. However, there are

artifacts associated with the generation of cDNA clones that may

have not been detected in our initial analyses such as internal

priming, priming from contaminating genomic DNA, retained introns

due to reverse transcription of unspliced precursor RNAs, and

reverse transcriptase errors that result in single base changes.

For further information about this sequence, including its location

and relationship to other sequences, please visit our Web site

(<http://fruitfly.berkeley.edu>) or send email to

cdna@fruitfly.berkeley.edu.

FEATURES

source

1. .1985

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

1. .1985

/gene="CG13127"

/db_xref="FLYBASE:FBgn0032176"

153. .1523

/gene="CG13127"

/note="Longest ORF"

/codon_start=1

/product="AT17414p"

/protein_id="AAM50552.1"

/db_xref="GI:21429748"

/db_xref="FLYBASE:FBgn0032176"

/translation="MLDVVDELKARRKHILNDIENNKVASNRMRKQLGITSGPVKL

AKASTARKKEPNILIGLNEENPLRSHYAHLSNIRVLMKGNESHRCAGDCGEV

LDVDCMGSAETWKPQNVKSFIEKPLPTSSPKKLADLIPVCSRTLRSLKSSPK

APAESFTNRSIGGQLDYHKIFDPLPLVEPMPEWDEWDLQKPVIERCIDALG

NDKSKELLRNKSNPVVSAANAKNFRFTKVAENAKNQDDKH1KQNGDILIDPDD

DEQIESLLKALRVLRGRQYVLA1LSNAHMPVLLDVAADRYGSKSCAGMKSIYKT

SFRIYKVEKERNKNMELKRSVTGLGSSISYASHKFKFSQVQVQRAEYNKLINE

LALEQSRUTLALRGYHLGGSIKDTFFAYMPARPDQIKRQHWKSYQFLDMVNFRLR

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 139 Row: i Column: 3
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Source

Location/Qualifiers

1..3430

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5732786"

/tissue_type="Brain, hippocampus"

/clone_lib="NIH_MGC_124"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

<1..3430

/gene="KIAA0602"

/db_xref="LocusID:23241"

<1..2787

/gene="KIAA0602"

/codon_start=1

/product="KIAA0602 protein"

/protein_id="AAH65220.1"

/db_xref="GI:40850908"

/db_xref="LocusID:23241"

/translation="PPGSHVTAPPPSARPARPPVRGPAAAPRRGGGAGAMAERG
 RGLPGAPALNTPVPMNLFWATVEVDGSSPCVPRICSLTKLVVFKLEKELISVY
 IAVKQSGKRILRSHIEIPLPPSQVETDLATFSLQYPHFLKREGNKQIMLQRRKRY
 KNRTILGYTLAAGISMAEVMOHPSGGQVLSLCSSKEAPVKAEEIWIASLSQPI
 DHEDSTMOAGPKAKSTDNYSEBEYFSFSEQASDDAVQGDLDDEDDFVGPKPKORR
 SIVRTSMTRQNFQKQKVALLRRFKVSDVLDSDQDPAEHIPEAEEDLDLLYDTLDM
 EHPSSGPMDEDDSVLSTPKPLRPYFGLSHSSQTEIGSIHSARSHKEPPSPADY
 PEKTRSLGQRQSDSVSDTVALGVPGPREHPGQPEDPEAEASTLDVFTERLPPSGRI
 TKTESLIVIPSTRSEKGRGRSTSLKERAQPNRANSLDNERCPTDARSQIQIP
 RKTVDQLNHLISDDLPENILVNTSDQGFSLDVLQRHTLPVCTCSPADYQAA
 ESTIVSRIQRYCNKSOPPTPVKIAVAGAHVLSAILRLFEQLSHKTPDWLGYMRF
 VIPLGSHPVARYLGSVDYRYNNPFQDLAWRDLFNKLEAQSVDPTDIVSRITQYIAG
 ANCAHQLPAEAWLITYKQSPDEESSQKFIFFGVVKGIVGIVESSATSGSDDDAAPSG
 SGTLSSTPPSPAPAAKEASPTPPSPSPSGSLSPSQGVGAELMGLQVDYWTAAQPAD
 RKRDAAKKDLPVTKTLKTPRSLQVSRLPSPSGEAAATPTMSMTVVTKEKNKVMFLP
 KKAKDKVSKSQCTEGISGLICTARQQQNMULRVLIDGVECSDVKFFQLAAQWSSHVK
 HFPICIFGHSKATP"

ORIGIN

Query Match 5.1%; Score 56; DB 9; Length 3430;
 Best Local Similarity 100.0%; Pred. No. 1.2e-18;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1034 ATAGCAAA 1089
 |||
 Db 3316 ATAGCAAA 3371
 |||

Search completed: February 6, 2005, 20:16:32
 Job time : 5152 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 15:46:16 ; Search time 226 Seconds
(without alignments)
7884.537 Million cell updates/sec

Title: US-10-088-830-1
Perfect score: 1089
Sequence: 1 gaaatcgagcagcgcaaa.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 1202784 seqs, 818138359 residues

Word size : 0

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/prodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/58_COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/68_COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	5.1	491	US-09-311-021-191	Sequence 191, Appl
2	54	5.0	176	US-09-621-976-13903	Sequence 13903, A
3	54	5.0	630	US-08-185-414E-1	Sequence 1, Appli
4	54	5.0	985	US-09-322-409-25	Sequence 25, Appl
5	54	5.0	985	US-09-322-409-27	Sequence 27, Appl
6	54	5.0	985	US-09-451-527-25	Sequence 25, Appl
7	54	5.0	985	US-09-451-527-27	Sequence 27, Appl
8	54	5.0	1023	US-08-252-966B-16	Sequence 16, Appl
9	54	5.0	1050	US-09-482-273-58	Sequence 58, Appl
10	54	5.0	1358	US-09-949-016-463	Sequence 463, App
11	54	5.0	1459	US-09-537-654-3	Sequence 3, Appli
12	54	5.0	1810	US-09-800-729-73	Sequence 73, Appl
13	54	5.0	1811	US-09-800-729-77	Sequence 77, Appl
14	54	5.0	2406	US-09-594-506-37	Sequence 37, Appl
15	54	5.0	3438	US-10-164-595-29	Sequence 29, Appl
16	53	4.9	150	US-09-621-976-8656	Sequence 8656, Ap
17	53	4.9	165	US-09-621-976-8127	Sequence 8127, Ap
18	53	4.9	194	US-09-621-976-801	Sequence 801, App
19	53	4.9	233	US-09-621-976-10675	Sequence 10675, A
20	53	4.9	249	US-09-621-976-19144	Sequence 19144, A
21	53	4.9	674	US-09-620-405B-465	Sequence 465, App
22	53	4.9	674	US-09-433-826B-465	Sequence 465, App
23	53	4.9	674	US-09-604-287A-465	Sequence 465, App
24	53	4.9	674	US-09-834-759-465	Sequence 465, App
25	53	4.9	674	US-09-590-751A-465	Sequence 465, App
26	53	4.9	674	US-09-551-621-465	Sequence 465, App
27	53	4.9	708	US-09-270-767-13081	Sequence 13081, A

28	53	4.9	857	1	US-08-308-883-1	Sequence 1, Appli
29	53	4.9	857	1	US-08-730-163-1	Sequence 1, Appli
30	53	4.9	857	3	US-08-256-799-1	Sequence 1, Appli
31	53	4.9	857	3	US-08-462-437-1	Sequence 1, Appli
32	53	4.9	941	4	US-09-205-258-186	Sequence 186, App
33	53	4.9	1066	1	US-08-157-101A-4	Sequence 4, Appli
34	53	4.9	1181	3	US-09-149-476-310	Sequence 310, App
35	53	4.9	1196	4	US-09-065-040-2	Sequence 2, Appli
36	53	4.9	1212	3	US-09-149-476-186	Sequence 186, App
37	53	4.9	1307	2	US-08-960-022-17	Sequence 17, Appl
38	53	4.9	1359	3	US-09-387-574-11	Sequence 11, Appl
39	53	4.9	1359	3	US-09-668-096-11	Sequence 11, Appl
40	53	4.9	1525	4	US-09-461-325-110	Sequence 110, App
41	53	4.9	1525	4	US-10-012-542-110	Sequence 110, App
42	53	4.9	1525	4	US-10-115-123-110	Sequence 110, App
43	53	4.9	1534	1	US-08-300-903A-6	Sequence 6, Appli
44	53	4.9	1534	4	US-08-988-197-6	Sequence 6, Appli
45	53	4.9	1534	4	US-10-385-072-6	Sequence 6, Appli

ALIGNMENTS

RESULT 1
US-09-311-021-191
; Sequence 191, Application US/093111021
; Patent No. 6706869
; GENERAL INFORMATION:
; APPLICANT: Wong, Gordon G.
; APPLICANT: Clark, Hilary
; APPLICANT: Fechtel, Kim
; APPLICANT: Agostino, Michael J. Inc.
; APPLICANT: Genetics Institute, Inc.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES ENCODING THEM
; FILE REFERENCE: GI 6300-11A
; CURRENT APPLICATION NUMBER: US/09/311,021
; CURRENT FILING DATE: 1999-05-13
; NUMBER OF SEQ ID NOS: 268
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 191
; LENGTH: 491
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-311-021-191

Query Match 5.1%; Score 55; DB 4; Length 491;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1035	TAGCAAA	1089
DB	430	TAGCAAA	484

RESULT 2
US-09-621-976-13903
; Sequence 13903, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 13903
; LENGTH: 176
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-13903

```
Query Match          5.0%; Score 54; DB 4; Length 176;
Best Local Similarity 100.0%; Pred. No. 3.4e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 85 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 138

RESULT 3
US-08-185-414E-1
; Sequence 1, Application US/08185414E
; Patent No. 5556953
; GENERAL INFORMATION:
; APPLICANT: Zhang, Lei
; APPLICANT: Vijay, Hari M.
; APPLICANT: Rode, Harold
; TITLE OF INVENTION: ALLERGEN OF CLADOSPORIUM HERBARUM
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: George A. Seaby
; ADDRESSEE: Seaby and Maclean
; STREET: 880 Wellington Street, Suite 708
; CITY: Ottawa
; COUNTRY: Canada
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5 inch diskette
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII Text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/185,414E
; FILING DATE: January 24, 1994
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: George A. Seaby
; REGISTRATION NUMBER: 24,034
; REFERENCE/DOCKET NUMBER: 1747
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (613) 232-5815
; TELEFAX: (613) 232-5831
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 630 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA and its encoded protein
; FEATURE: Open reading frame of CH2.1
US-08-185-414E-1

Query Match          5.0%; Score 54; DB 1; Length 630;
Best Local Similarity 100.0%; Pred. No. 3e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 544 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 597

RESULT 4
US-09-322-409-25
; Sequence 25, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C1
```

```
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (74)..(901)
US-09-322-409-25

Query Match          5.0%; Score 54; DB 4; Length 985;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 921 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 974

RESULT 5
US-09-322-409-27/c
; Sequence 27, Application US/09322409
; Patent No. 6471957
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C1
; CURRENT APPLICATION NUMBER: US/09/322,409
; CURRENT FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 154
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-322-409-27

Query Match          5.0%; Score 54; DB 4; Length 985;
Best Local Similarity 100.0%; Pred. No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 65 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 6
US-09-451-527-25
; Sequence 25, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Kee
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES, AND USES THEREOF
; FILE REFERENCE: IM-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
```



```

; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 25
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (74)..(901)
US-09-451-527-25

Query Match      5.0%; Score 54; DB 4; Length 985;
Best Local Similarity 100.0%; Pred.No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
          |||||||
Db       921 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 974

RESULT 7
US-09-451-527-27/c
; Sequence 27, Application US/09451527
; Patent No. 6482403
; GENERAL INFORMATION:
; APPLICANT: Sim, Gek-Keo
; APPLICANT: Yang, Shumin
; APPLICANT: Dreitz, Matthew J.
; APPLICANT: Wonderling, Ramani S.
; TITLE OF INVENTION: CANINE AND FELINE IMMUNOREGULATORY PROTEINS, NUCLEIC
; FILE REFERENCE: IW-2-C2
; CURRENT APPLICATION NUMBER: US/09/451,527
; CURRENT FILING DATE: 1999-12-01
; EARLIER APPLICATION NUMBER: 09/322,409
; EARLIER FILING DATE: 1999-05-28
; EARLIER APPLICATION NUMBER: 60/087,306
; EARLIER FILING DATE: 1998-05-29
; NUMBER OF SEQ ID NOS: 174
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 27
; LENGTH: 985
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-451-527-27

Query Match      5.0%; Score 54; DB 4; Length 985;
Best Local Similarity 100.0%; Pred.No. 2.9e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
          |||||||
Db       65  AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 8
US-08-252-966B-16
; Sequence 16, Application US/08252966B
; Patent No. 5624818
; GENERAL INFORMATION:
; APPLICANT: Eisenman, Robert N.
; APPLICANT: Hurlin, Peter J.
; APPLICANT: Ayer, Donald E.
; TITLE OF INVENTION: Regulatory Proteins that Dimerize with
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Christensen, O'Connor, Johnson, and KindnessP/LC
; STREET: 1420 Fifth Ave., Suite 2800
; CITY: Seattle
; STATE: Washington

```



```
; APPLICANT: Rafalski, J. Antoni
; TITLE OF INVENTION: Thiamine Biosynthetic Enzymes
; FILE REFERENCE: BB1372 US NA
; CURRENT APPLICATION NUMBER: US/09/594,506
; CURRENT FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 60/139,556
; PRIOR FILING DATE: 1999-06-16
; NUMBER OF SEQ ID NOS: 45
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 37
; LENGTH: 2406
; TYPE: DNA
; ORGANISM: Triticum aestivum
US-09-594-506-37

Query Match          5.0%; Score 54; DB 4; Length 2406;
Best Local Similarity 100.0%; Pred. No. 2.7e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
      |||||||
Db 2351 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2404
      |||||||

RESULT 15
US-10-164-595-29
; Sequence 29, Application US/10164595
; Patent No. 6657054
; GENERAL INFORMATION:
; APPLICANT: OriGene Technologies, Inc
; TITLE OF INVENTION: Regulated Angiogenesis Genes and Polypeptides
; FILE REFERENCE: 1U 103 R1
; CURRENT APPLICATION NUMBER: US/10/164,595
; CURRENT FILING DATE: 2002-06-10
; NUMBER OF SEQ ID NOS: 80
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 29
; LENGTH: 3438
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (527)..(2701)
; OTHER INFORMATION:
US-10-164-595-29

Query Match          5.0%; Score 54; DB 4; Length 3438;
Best Local Similarity 100.0%; Pred. No. 2.6e-13;
Matches 54; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1036 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
      |||||||
Db 3345 AGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3398
      |||||||

Search completed: February 6, 2005, 17:39:37
Job time : 228 secs
```

This Page Blank (uspto)


```

QY 75 CCGCGTGCACATCTTCGAGCTTCCAGTGTCCCGCGCTTCCCGAA 121
    |||||
Db 609 CCGCGTGCACATCTTCGAGCTTCCAGTGTCCCGCGCTTCCCGAA 655

RESULT 2
CD892664
LOCUS G118.121J23F010725 G118 Triticum aestivum cDNA clone G118121J23,
DEFINITION mRNA sequence.
ACCESSION CD892664
VERSION CD892664.1 GI:32663198
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 454)
Genoplante, a major partnership french program in plant genomics
REFERENCE 1
AUTHORS Genoplante,
TITLE Unpublished (2003)
JOURNAL Contact: Genoplante
COMMENT Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES
    source
    1..454
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="G118121J23"
    /tissue_type="grain (118 degrees per day after
    pollination)"
    /clone_lib="G118"

ORIGIN
    Query Match 9.0%; Score 98; DB 6; Length 454;
    Best Local Similarity 99.3%; Pred. No. 5.9e-36;
    Matches 148; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 689 AAGTTTGGCCATTTTCAGTTCAATGGTGCACCATTCACATTCGATGATCTCTCAATC 748
    |||||
Db 92 AAGTTTGGCCATTTTCAGTTCAATGGTGCACCATTCACATTCGATGATCTCTCAATC 151

QY 749 CTTGAGGGGTAAGCGGTAACAGCATAGAGAGAGTGGCGCGGCACCCCTTCACATAGAGA 808
    |||||
Db 152 CTTGAGGGGTAAGCGGTAACAGCATAGAGAGTGGCGCGGCACCCCTTCACATAGAGA 211

QY 809 CTCAGAAATATTACAAATCAATTAAGCT 837
    |||||
Db 212 CTCAGAAATATTACAAATCAATTAAGCT 240

RESULT 3
CD882927
LOCUS F1.111L02F010430 F1 Triticum aestivum cDNA clone F111L02, mRNA
DEFINITION sequence.
ACCESSION CD882927
VERSION CD882927.1 GI:32645392
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 649)
Genoplante, a major partnership french program in plant genomics
REFERENCE 1
AUTHORS Genoplante,
TITLE Unpublished (2003)
JOURNAL Contact: Genoplante
COMMENT Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES
    source
    1..454
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="G118121J23"
    /tissue_type="grain (118 degrees per day after
    pollination)"
    /clone_lib="G118"

ORIGIN
    Query Match 9.1%; Score 88; DB 6; Length 649;
    Best Local Similarity 100.0%; Pred. No. 4e-31;
    Matches 88; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 407 GCMAAAGAAAAAAGAGATACGGTGGATGGCGCTTCAAAATTACAGATATGAAAAATA 466
    |||||
Db 119 GCMAAAGAAAAAAGAGATACGGTGGATGGCGCTTCAAAATTACAGATATGAAAAATA 178

QY 467 AAGAAGCTTTGAGGAAGTTCTGTAAGAAC 494
    |||||
Db 179 AAGAAGCTTTGAGGAAGTTCTGTAAGAAC 206

RESULT 4
CA802821
LOCUS sau42502.y1 Gm-c1071 Glycine max cDNA clone soybean EST 01-JUL-2004
DEFINITION Gm-c1071-3748 5', mRNA sequence.
ACCESSION CA802821
VERSION CA802821.1 GI:26059907
KEYWORDS EST.
SOURCE Glycine max (soybean)
ORGANISM Glycine max
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids 1; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 301)
Shoenaker, R., Keim, P., Vodkin, L., Erpelting, J., Corvelli, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurr, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Putative full length read vector to vector length is 467
Seq primer: -40RP from Gibco
High quality sequence stop: 286.
Location/Qualifiers

FEATURES
    Location/Qualifiers
    1..649
    /organism="Triticum aestivum"
    /mol_type="mRNA"
    /cultivar="recital"
    /db_xref="taxon:4565"
    /clone="F111L02"
    /tissue_type="leaf one"
    /clone_lib="F1"

```

[illegible]

AL697897
 LOCUS DKFZP686A19105 169 bp mRNA linear EST 04-SEP-2003
 DEFINITION DKFZP686A19105_r1 686 (synonym: hlcc3) Homo sapiens cDNA clone
 VERSION AL697897
 KEYWORDS AL697897.1 GI:19618437
 EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 169)
 AUTHORS Koehrer, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.
 TITLE EST (Koehrer, K., Beyer, A., Mewes, H.W., Weil, B. and Wiemann, S.)
 JOURNAL Unpublished (1999)
 COMMENT Contact: MIPS
 Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
 This is the 5' sequence of the clone insert
 Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
 Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
 sequenced by BMFZ (Biomedical Research Center at the Charite,
 Berlin/Germany) within the cDNA sequencing consortium of the German
 Genome Project.
 No sl sequence available.
 This clone (DKFZP686A19105) is available at the RZPD in Berlin.
 Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
 Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.

FEATURES
 source
 1..169
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="DKFZP686A19105"
 /dev_stage="adult"
 /lab_host="DH10B"
 /clone_lib="686 (synonym: hlcc3)"
 /note="Vector: pTriplex2; Site_1: SfiIA; Site_2: SfiIB;
 cDNA-collection"

ORIGIN

Query Match 5.1%; Score 56; DB 1; Length 169;
 Best Local Similarity 100.0%; Pred. No. 1.3e-15;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1034 ATACCAA 1089
 Db 60 ATACCAA 115

RESULT 8
 BGI79244
 LOCUS BGI79244 177 bp mRNA linear EST 06-FEB-2001
 DEFINITION 602330836F1 NIH_MGC_91 Homo sapiens cDNA clone IMAGE:4432161 5',
 mRNA sequence.
 ACCESSION BGI79244
 VERSION BGI79244.1 GI:12685876
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 177)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF

cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 cDNA distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLAM10188 row: g column: 10
 High quality sequence start: 3
 High quality sequence stop: 169.

FEATURES
 source

1..177
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4432161"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_91"
 /note="Organ: prostate; Vector: pCMV-SPORT6; Site_1: NotI;
 Site_2: SalI; Cloned unidirectionally; oligo-dr primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 5.1%; Score 56; DB 4; Length 177;
 Best Local Similarity 100.0%; Pred. No. 1.3e-15;
 Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 Qy 1034 ATACCAA 1089
 Db 98 ATACCAA 153

RESULT 9

CO249876
 LOCUS CO249876 214 bp mRNA linear EST 23-JUN-2004
 DEFINITION AGENCOURT 26529183 NIH_MGC_212 Homo sapiens cDNA clone
 IMAGE:30925067 5', mRNA sequence.

ACCESSION CO249876
 VERSION CO249876.1 GI:49121395
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 214)
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute / NIH
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-remail.nih.gov

Tissue Procurement: Mary Hendrix
 cDNA Library Preparation: M. Bento Soares, University of Iowa
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 cDNA distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDAM1168 row: i column: 12
 High quality sequence stop: 214.

FEATURES
 source

1..214
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:30925067"
 /tissue_type="Chondrosarcoma Lung Metastasis cell lines"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH_MGC_212"
 /note="Organ: Lung; Vector: pYX-Asc; Site_1: EcoR I;
 Site_2: Not I; The library was constructed according
 Ronaldo, Lennon and Soares, Genome Research, 6:791-806,
 1996. Denatured RNA was size fractionated on a 1% agarose


```

RESULT 12
AW734913
LOCUS
DEFINITION AW734913 269 bp mRNA linear EST 14-JUL-2004
Gm-c1016-10054 5', mRNA sequence.
ACCESSION
VERSION AW734913
KEYWORDS
SOURCE EST.
ORGANISM Glycine max (soybean)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 269)
REFERENCE
AUTHORS Shoemaker,R., Keim,P., Vodkin,L., Erpelting,J., Coryell,V.,
Khanna,A., Bolla,B., Marra,M., Hillier,L., Kucaba,T., Martin,J.,
Beck,C., Wylie,T., Underwood,K., Steptoe,M., Theising,B., Allen,M.,
Bowers,Y., Person,B., Swaller,T., Gibbons,M., Pape,D., Harvey,N.,
Schurk,R., Ritter,E., Kohn,S., Shin,T., Jackson,K., Cardenas,M.,
McCann,R., Waterston,R. and Wilson,R.
JOURNAL Public Soybean EST Project
COMMENT Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
When it has been determined, an EST from the other end of this
clone is listed in the 'Other ESTs on clone' field. This clone is
available through: Biogenetic Services, 801 32nd Ave. Brookings, SD
57006 USA (phone: 800 423 4163; email: info@biogeneticservices.com)
Seq primer: -40RP from Gibco
High quality sequence stop: 227.
FEATURES
source
1..269
Location/Qualifiers
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-10054"
/tissue_type="immature flowers of field grown plants"
/lab_host="XL10-Gold"
/clone_lib="Gm-c1016"
/notes="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pBluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adapters
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pBluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelting."
ORIGIN
Query Match 5.1%; Score 56; DB 2; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1034 ATACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 207 ATACCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 262
RESULT 13
AL046990
LOCUS
DEFINITION AL046990 287 bp mRNA linear EST 04-SEP-2003
DXFZ586L2017 r1.586 (synonym: hutel) Homo sapiens cDNA clone
DXFZ586L2017 5', mRNA sequence.
ACCESSION
VERSION AL046990
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 287)
REFERENCE
AUTHORS Koehler,K., Beyer,A., Mewes,H.W., Gassenhuber,J. and Wiemann,S.
JOURNAL EST (Koehler, et al.)
COMMENT Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5435046.
Contact: MIPS
MIPS
Ingolstaedter Landstr.1, D-85764 Neuherberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
No sl sequence available.
This clone (DXFZ586L2017) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
FEATURES
source
1..287
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DXFZ586L2017"
/tissue_type="uterus"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="586 (synonym: hutel)"
/notes="Vector: pSport1; Site_1: NotI; Site_2: SalI/MluI"
ORIGIN
Query Match 5.1%; Score 56; DB 1; Length 287;
Best Local Similarity 100.0%; Pred. No. 1.3e-15;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1034 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 191 ATAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 246
RESULT 14
BI756116
LOCUS
DEFINITION BI756116 290 bp mRNA linear EST 25-SEP-2001
603030013F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:5200226 5',
mRNA sequence.
ACCESSION
VERSION BI756116
KEYWORDS
SOURCE EST.
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 290)
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@email.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be

```

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LHAM1501 row: j column: 03

High quality sequence stop: 236.

FEATURES

source

1. 290

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5200226"

/lab_host="DH10B"

/clone_lib="NIH_MGC_114"

/note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 5.1%; Score 56; DB 4; Length 290;

Best Local Similarity 100.0%; Pred. No. 1.3e-15;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1034 ATAGCAA 1089

Db 138 ATAGCAA 193

RESULT 15

AA413331/c

LOCUS

AA413331 294 bp mRNA linear EST 02-MAY-1997

DEFINITION AGEST00036 Anopheles gambiae adult pSport cDNA Anopheles gambiae

CDNA clone cc38 3', mRNA sequence.

ACCESSION

AA413331

VERSION

AA413331.1

GI:2071917

EST.

SOURCE

ORGANISM Anopheles gambiae (African malaria mosquito)

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea;

Anopheles.

1 (bases 1 to 294)

Cornel, A.J., Kumar, V., Mukabayire, O., Salazar Rafferty, C.,

Petrarca, V., Coluzzi, M. and Collins, F.H.

A comprehensive physical map of the malaria vector Anopheles

gambiae

Unpublished (1997)

Other ESTs: AGEST00035

Contact: Salazar Rafferty, C.; and Collins, FH

Vector Genetics Section

Centers for Disease Control and Prevention

MS F-22, 4770 Buford Hwy, Chamblee, GA 30341

Tel: 770 488 7463

Fax: 770 488 7469

Email: czs7@cdc.gov

Mapping location: 34C in 3R chromosome

Seq primer: M13 Universal

High quality sequence stop: 294.

FEATURES

source

1. 294

/organism="Anopheles gambiae"

/mol_type="mRNA"

/strain="G3"

/db_xref="taxon:7165"

/clone="cc38"

/sex="Male and Female"

/clone_lib="Anopheles gambiae adult pSport cDNA"

/note="Site 1: Sali; Site 2: NotI; See: Salazar, C.E., et

al. Insect Molecular Biology (1994), 3:1-13."

ORIGIN

Query Match 5.1%; Score 56; DB 1; Length 294;

Best Local Similarity 100.0%; Pred. No. 1.3e-15;

Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1034 ATAGCAA 1089

Db 66 ATAGCAA 11

Search completed: February 6, 2005, 18:50:42

Job time : 4259 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 14:35:16 ; Search time 698 Seconds
(without alignments)
9235.814 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcggcagcagcgca.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: OLIGO_NUC
Gapop_60.0 , Gapext 60.0

Searched: 4390206 seqs, 2959870667 residues

Word size : 0

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004as:*
- 13: Geneseqn2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the chance being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	965	88.6	1089	4	Aaf80144 Nucleotid
2	168	15.4	168	4	Aaf80149 DNA fragm
3	168	15.4	168	4	Aaf80148 DNA fragm
4	153	14.0	153	4	Aaf80147 DNA fragm
5	150	13.8	150	4	Aaf80150 DNA fragm
6	114	10.5	156	4	Aaf80146 DNA fragm
7	56	5.1	668	5	Aah64916 Human sec
8	55	5.1	376	5	Abv58410 Human pro
9	55	5.1	396	3	Aad02077 cDNA enco
10	55	5.1	410	8	Abx42327 Bovine ES
11	55	5.1	443	4	Aai90926 Human pol
12	55	5.1	491	3	Aaz52570 Human sec
13	55	5.1	917	3	Aac77635 Human can
14	55	5.1	2435	12	Adq22529 Human sof
15	55	5.1	3759	3	Aac58600 Human PRO
16	55	5.1	3819	2	Aaz34292 Human PRO
17	55	5.1	3819	3	Aac78580 Human PRO
18	55	5.1	3819	4	Aas21446 Human CDN
19	55	5.1	3819	6	Abk33602 cDNA enco
20	55	5.1	3819	6	AbL88117 Human PRO

21	55	5.1	3819	6	ABL95606	Human ang
22	55	5.1	3819	8	ACA63860	Novel hum
23	55	5.1	3819	8	ACA03805	cDNA enco
24	55	5.1	3819	8	ACA72024	Human sec
25	55	5.1	3819	8	ABX89343	DNA enco
26	55	5.1	3819	8	ABX92664	cDNA enco
27	55	5.1	3819	8	ACD41997	Human sec
28	55	5.1	3819	8	ACA66405	Human CDN
29	55	5.1	3819	8	ACA68563	Novel hum
30	55	5.1	3819	8	ACA04226	Human CDN
31	55	5.1	3819	9	ADA45924	Novel hum
32	55	5.1	3819	9	ADA76355	Human PRO
33	55	5.1	3819	9	ABT44292	Human PRO
34	55	5.1	3819	9	ADA19005	Human PRO
35	55	5.1	3819	9	ADA61628	Homo sapi
36	55	5.1	3819	9	ADB19413	Novel hum
37	55	5.1	3819	9	ADB27954	cDNA enco
38	55	5.1	3819	9	ADA86433	Novel hum
39	55	5.1	3819	9	ADB15997	Human PRO
40	55	5.1	3819	9	ADA47783	Human PRO
41	55	5.1	3819	9	ADA67578	Human PRO
42	55	5.1	3819	9	ADB30585	cDNA enco
43	55	5.1	3819	9	ADA85881	Novel hum
44	55	5.1	3819	9	ADA97093	Human PRO
45	55	5.1	3819	9	ADA79397	Human PRO

ALIGNMENTS

RESULT 1

AAF80144

ID AAF80144 standard; DNA; 1089 BP.

AC AAF80144;

DT 11-JUN-2001; (first entry)

DE Nucleotide sequence of a wheat E2F-dimerisation partner (DP) protein.

KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;

KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;

OS Triticum monococcum.

FT Key Location/Qualifiers

FT CDS 20..805

FT /*tag= a

FT /product= "E2F-dimerisation partner (DP) protein"

PN WO200121644-A2.

XX 29-MAR-2001.

XX 25-SEP-2000; 2000WO-EP009325.

XX 24-SEP-1999; 99ES-00002127.

XX 11-NOV-1999; 99ES-00002474.

XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

XX Gutierrez-Armenta C, Ramirez-Parra E;

XX WPI; 2001-257972/26.

XX P-PSDB; AAB67762.

XX New isolated, enriched, cell free and/or recombinant nucleic acid useful for e.g. altering cell proliferation characteristic such as to alter plant cell, organ or tissue size.

XX Claim 13; Fig 1; 77pp; English.

CC The present sequence encodes a E2F-dimerisation partner (DP) protein. The
 CC proteins acts as a plant E2F transcription factor. E2F and DP are two
 CC proteins that hetero-dimerise to form an active transcription factor that
 CC regulates G1 to S phase of the cell cycle, and later, the expression of
 CC genes required for S-phase progression. E2F and retinoblastoma protein
 CC also interact as a hetero-dimer in cells to suppress certain genes. This
 CC repression involves binding of the retinoblastoma protein to the E2F-DP
 CC dimer that is in turn bound to sites on DNA through the E2F DNA binding
 CC domain. DP proteins can be modulated to alter plant cell, organ or tissue
 CC shape, and particularly to alter cell proliferation characteristic such
 CC as to alter plant cell, organ or tissue size
 XX
 SQ Sequence 1089 BP; 367 A; 215 C; 261 G; 246 T; 0 U; 0 Other;

Query Match 88.6%; Score 965; DB 4; Length 1089;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 965; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 125 GCGGTAATGCGTCCAAAGGAAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGGAG 184
 DB 125 GCGGTAATGCGTCCAAAGGAAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGGAG 184

QY 185 AAGGCTGCGGACCGAGGATCACCGGTTGGGGGCTCGCGAGTACAGCAAAATAGTTTGT 244
 DB 185 AAGGCTGCGGACCGAGGATCACCGGTTGGGGGCTCGCGAGTACAGCAAAATAGTTTGT 244

QY 245 GAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGGTTCCAGCAGCAAAATTTATCA 304
 DB 245 GAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGGTTCCAGCAGCAAAATTTATCA 304

QY 305 GAGCTGAAGTCCATGCGACATATTTGTCAGGGTTTGTGAGAGATATTTAGCGCGAGA 364
 DB 305 GAGCTGAAGTCCATGCGACATATTTGTCAGGGTTTGTGAGAGATATTTAGCGCGAGA 364

QY 365 GTGTATGATGCTTTCAACGTTCTCATTTGCACTTCGTTTATTCGCAAAAGAAAAAGGAG 424
 DB 365 GTGTATGATGCTTTCAACGTTCTCATTTGCACTTCGTTTATTCGCAAAAGAAAAAGGAG 424

QY 425 ATACGGTGAATGGGCTTTCAAAATTCAGATATGAAAGCTTTGAGGAAGTT 484
 DB 425 ATACGGTGAATGGGCTTTCAAAATTCAGATATGAAAGCTTTGAGGAAGTT 484

QY 485 CGTAAAGAACTCGTCAACAGATTTAGGAACAAGAGCACTCTCCAGGAATTCGAAAAA 544
 DB 485 CGTAAAGAACTCGTCAACAGATTTAGGAACAAGAGCACTCTCCAGGAATTCGAAAAA 544

QY 545 CAGTTTATGATGATCTCCAAACATCAAGTTACGTAAACAACTGGAAGCTCAGCAGAG 604
 DB 545 CAGTTTATGATGATCTCCAAACATCAAGTTACGTAAACAACTGGAAGCTCAGCAGAG 604

QY 605 AATGTTAATGGCATCGGCTTCATTCGTTATGTTGGTCAAGACATCTAGGAAGCAAGGGTG 664
 DB 605 AATGTTAATGGCATCGGCTTCATTCGTTATGTTGGTCAAGACATCTAGGAAGCAAGGGTG 664

QY 665 GAAATTTGAGATTTTCAGATGACTCGAAGTTTGGCCATTCGAGTTCAATGTTGACCAATTC 724
 DB 665 GAAATTTGAGATTTTCAGATGACTCGAAGTTTGGCCATTCGAGTTCAATGTTGACCAATTC 724

QY 725 ACATTTGATGATGATCTCTCAATTCCTTGGGGGTTAAGGCGTAAACAGCATAGGAAGCT 784
 DB 725 ACATTTGATGATGATCTCTCAATTCCTTGGGGGTTAAGGCGTAAACAGCATAGGAAGCT 784

QY 785 GGGCGGCGCAACCTTCATCTAGAGCTCAAGAATATTACAAATTAAGAAAGTTTAGAA 844
 DB 785 GGGCGGCGCAACCTTCATCTAGAGCTCAAGAATATTACAAATTAAGAAAGTTTAGAA 844

QY 845 CTGGCACAGCCGATCTTTTGGCACAGCTATGATAGCTATATCTCATGAAACTTG 904
 DB 845 CTGGCACAGCCGATCTTTTGGCACAGCTATGATAGCTATATCTCATGAAACTTG 904

QY 905 ACCTAGTTTATAGGACAGTCTCTCAGGCTTGAAGATTTTAACTGCAAAATTTTCTC 964
 DB 905 ACCTAGTTTATAGGACAGTCTCTCAGGCTTGAAGATTTTAACTGCAAAATTTTCTC 964

QY 965 CTTTTTGTGCTAGCAGGTTATTAGTCTCAGATAGATGATTCATATATGCTGCTATG 1024
 DB 965 CTTTTTGTGCTAGCAGGTTATTAGTCTCAGATAGATGATTCATATATGCTGCTATG 1024

QY 1025 AAAACATTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084
 DB 1025 AAAACATTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1084

QY 1085 AAAAA 1089
 DB 1085 AAAAA 1089

RESULT 2
 AAF80149
 ID AAF80149 standard; DNA; 168 BP.
 XX
 AC AAF80149;
 XX
 DT 11-JUN-2001 (first entry)
 XX
 DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
 XX
 KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
 KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
 KW ss.
 XX
 OS Triticum monococcum.
 XX
 PN WO200121644-A2.
 XX
 PD 29-MAR-2001.
 XX
 PF 25-SEP-2000; 2000WO-EP009325.
 XX
 PR 24-SEP-1999; 99ES-00002127.
 PR 11-NOV-1999; 99ES-00002474.
 XX
 PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 XX
 PI Gutierrez-Armenta C, Ramirez-Parra E;
 XX
 DR WPI; 2001-257972/26.
 DR P-PSDB; AAB67767.
 XX
 PT New isolated, enriched, cell free and/or recombinant nucleic acid useful
 PT for e.g. altering cell proliferation characteristic such as to alter
 PT plant cell, organ or tissue size.
 XX
 PS Disclosure; Page 74-75; 77pp; English.
 XX
 CC AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
 CC partner (DP) protein. The protein acts as a plant E2F transcription
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an
 CC active transcription factor that regulates G1 to S phase of the cell
 CC cycle, and later, the expression of genes required for S-phase
 CC progression. E2F and retinoblastoma protein also interact as a hetero-
 CC dimer in cells to suppress certain genes. This repression involves
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
 CC be modulated to alter plant cell, organ or tissue shape, and particularly
 CC to alter cell proliferation characteristic such as to alter plant cell,
 CC organ or tissue size
 XX
 SQ Sequence 168 BP; 62 A; 38 C; 33 G; 35 T; 0 U; 0 Other;

Query Match 15.4%; Score 168; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 1.8e-49;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 485 CGTAAAGAACTCGTCAACAGATTTAGGAACAAGGCACTCTCCAGGAATTCGAAAAA 544
 DB 485 CGTAAAGAACTCGTCAACAGATTTAGGAACAAGGCACTCTCCAGGAATTCGAAAAA 544

```
Db 1 CGTAAAGAACTCGTCAACAGATTAGGAAACAAGAGGCACCTCTCCAGGAAATCGAAAAA 60
QY 545 CAGTTTGATGATCTCCAAACATCAAGTTACGTAAACCAACACTGGAAGCTCAGCAGAG 604
Db 61 CAGTTTGATGATCTCCAAACATCAAGTTACGTAAACCAACACTGGAAGCTCAGCAGAG 120
QY 605 AATGTTAATGGCATCGCCCTTCATTCCGTTATGGTCAAGACATCTAGG 652
Db 121 AATGTTAATGGCATCGCCCTTCATTCCGTTATGGTCAAGACATCTAGG 168

RESULT 3
AAF80148
ID AAF80148 standard; DNA; 168 BP.
XX AC
XX AAF80148;
XX DT 11-JUN-2001 (first entry)
XX DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
XX KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
XX KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
XX KW ss.
XX OS Triticum monococcum.
XX PN WO200121644-A2.
XX PD 29-MAR-2001.
XX PF 25-SEP-2000; 2000WO-EP009325.
XX PR 24-SEP-1999; 99ES-00002127.
XX PR 11-NOV-1999; 99ES-00002474.
XX PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX PI Gutierrez-Armenta C, Ramirez-Parra E;
XX WPI; 2001-257972/26.
XX DR P-PSDB; AAB67766.
XX PT New isolated, enriched, cell free and/or recombinant nucleic acid useful
XX PT for e.g. altering cell proliferation characteristic such as to alter
XX PT plant cell, organ or tissue size.
XX PS Disclosure; Page 73-74; 77pp; English.
XX CC AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
XX CC partner (DP) protein. The protein acts as a plant E2F transcription
XX CC factor. E2F and DP are two proteins that hetero-dimerise to form an
XX CC active transcription factor that regulates G1 to S phase of the cell
XX CC cycle, and later, the expression of genes required for S-phase
XX CC progression. E2F and retinoblastoma protein also interact as a hetero-
XX CC dimer in cells to suppress certain genes. This repression involves
XX CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
XX CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
XX CC be modulated to alter plant cell, organ or tissue shape, and particularly
XX CC to alter cell proliferation characteristic such as to alter plant cell,
XX CC organ or tissue size
XX SQ Sequence 168 BP; 59 A; 19 C; 44 G; 46 T; 0 U; 0 Other;

Query Match 15.4%; Score 168; DB 4; Length 168;
Best Local Similarity 100.0%; Pred. No. 1.8e-49;
Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 317 ATGGCACATATTGGTCAAGGTTTGATGAGAGAATATTAGCGGAGAGTGATGATGCT 376
Db 1 ATGGCACATATTGGTCAAGGTTTGATGAGAGAATATTAGCGGAGAGTGATGATGCT 60
QY 377 TTCAACGTTCTCATTCGACTTCGTGTTATTGCAAAAAGAAAAAGAGATACGGTGGATG 436
```

```
Db 61 TTCAACGTTCTCATTCGACTTCGTGTTATTGCAAAAAGAAAAAGAGATACGGTGGATG 120
QY 437 GGCCTTTTCAAAATACAGATATGAAAAAATAAAGAGCTTTGAGGAAGTT 484
Db 121 GGCCTTTTCAAAATACAGATATGAAAAAATAAAGAGCTTTGAGGAAGTT 168

RESULT 4
AAF80147
ID AAF80147 standard; DNA; 153 BP.
XX AC
XX AAF80147;
XX DT 11-JUN-2001 (first entry)
XX DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
XX KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
XX KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
XX KW ss.
XX OS Triticum monococcum.
XX PN WO200121644-A2.
XX PD 29-MAR-2001.
XX PF 25-SEP-2000; 2000WO-EP009325.
XX PR 24-SEP-1999; 99ES-00002127.
XX PR 11-NOV-1999; 99ES-00002474.
XX PA (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
XX PI Gutierrez-Armenta C, Ramirez-Parra E;
XX WPI; 2001-257972/26.
XX DR P-PSDB; AAB67765.
XX PT New isolated, enriched, cell free and/or recombinant nucleic acid useful
XX PT for e.g. altering cell proliferation characteristic such as to alter
XX PT plant cell, organ or tissue size.
XX PS Disclosure; Page 72-73; 77pp; English.
XX CC AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
XX CC partner (DP) protein. The protein acts as a plant E2F transcription
XX CC factor. E2F and DP are two proteins that hetero-dimerise to form an
XX CC active transcription factor that regulates G1 to S phase of the cell
XX CC cycle, and later, the expression of genes required for S-phase
XX CC progression. E2F and retinoblastoma protein also interact as a hetero-
XX CC dimer in cells to suppress certain genes. This repression involves
XX CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
XX CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
XX CC be modulated to alter plant cell, organ or tissue shape, and particularly
XX CC to alter cell proliferation characteristic such as to alter plant cell,
XX CC organ or tissue size
XX SQ Sequence 153 BP; 55 A; 25 C; 47 G; 26 T; 0 U; 0 Other;

Query Match 14.0%; Score 153; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 3.4e-44;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 164 GATAAGATAGGAAGAAGAGAGGCTCGGCACCGAGGATCACCGTTGGGGGCTCCGC 223
Db 1 GATAAGATAGGAAGAAGAGAGGCTCGGCACCGAGGATCACCGTTGGGGGCTCCGC 60
QY 224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGACACATCATGAG 283
Db 61 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGACACATCATGAG 120
```

QY	284	GTTCGACAGCAAAATTTATTTCAGAGCTGAAGTCC	316
Db	121	GTTCGACAGCAAAATTTATTTCAGAGCTGAAGTCC	153
RESULT 5			
AAF80150		AAF80150 standard; DNA; 150 BP.	
XX	AC	AC	XX
XX	11-JUN-2001	(first entry)	XX
XX		DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.	XX
DE	XX	E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;	XX
KW	KW	S phase; cell cycle; retinoblastoma protein; alter cell proliferation;	XX
KW	KW	ss.	XX
OS	XX	Triticum monoccoccum.	XX
XX	WO200121644-A2.		XX
XX	29-MAR-2001.		XX
XX	25-SEP-2000; 2000WO-EP009325.		XX
PF	XX	24-SEP-1999; 99ES-00002127.	XX
PR	XX	11-NOV-1999; 99ES-00002474.	XX
XX		(CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.	XX
XX		Gutierrez-Armenta C, Ramirez-Parra E;	XX
PI	XX	WPI; 2001-257972/26.	XX
DR	XX	P-PSDB; AAB67768.	XX
PT	XX	New isolated, enriched, cell free and/or recombinant nucleic acid useful	XX
PT	XX	for e.g. altering cell proliferation characteristic such as to alter	XX
PT	XX	plant cell, organ or tissue size.	XX
XX		Disclosure; Page 75-76; 77pp; English.	XX
PS	XX	AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation	XX
CC	XX	partner (Dp) protein. The protein acts as a plant E2F transcription	XX
CC	XX	factor. E2F and DP are two proteins that hetero-dimerise to form an	XX
CC	XX	active transcription factor that regulates G1 to S phase of the cell	XX
CC	XX	cycle, and later, the expression of genes required for S-phase	XX
CC	XX	progression. E2F and retinoblastoma protein also interact as a hetero-	XX
CC	XX	dimer in cells to suppress certain genes. This repression involves	XX
CC	XX	binding of the retinoblastoma protein to the E2F-Dp dimer that is in turn	XX
CC	XX	bound to sites on DNA through the E2F DNA binding domain. DP proteins can	XX
CC	XX	be modulated to alter plant cell, organ or tissue shape, and particularly	XX
CC	XX	to alter cell proliferation characteristic such as to alter plant cell,	XX
XX	XX	organ or tissue size	XX
XX	XX	Sequence 150 BP; 41 A; 34 C; 39 G; 36 T; 0 U; 0 Other;	XX
Query March 13.8%; Score 150; DB 4; Length 150;			
Best Local Similarity 100.0%; Pred. No. 3.8e-43;			
Matches 150; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
QY	653	AAAGCAAGGGTGGAATTCAGATTCAGATTCGATCGAAGTTGGCCCAATTCGAGTTCAAT	712
Db	1	AAAGCAAGGGTGGAATTCAGATTCAGATTCGATCGAAGTTGGCCCAATTCGAGTTCAAT	60
QY	713	GGTCGACCAATTCACATTCGATGATGATCTCTCAATCCTTGAGGGGGTAAGCGGTAAACAGC	772
Db	61	GGTCGACCAATTCACATTCGATGATGATCTCTCAATCCTTGAGGGGGTAAGCGGTAAACAGC	120
QY	773	ATAGGAAGAGCTGGCGCGCCACCCCTTCAC	802
Db	131	ATAGGAAGAGCTGGCGCGCCACCCCTTCAC	150

XX 11-SEP-2001 (first entry)
XX Human secreted protein cDNA, SEQ ID NO: 192.
XX
XX Human; secreted protein; gene therapy; vaccine; treatment; diagnosis;
KW GENSET; ss.
XX
XX Homo sapiens.
XX
XX WO200142451-A2.
XX
XX 14-JUN-2001.
XX
XX 07-DEC-2000; 2000WO-1B001938.
XX
XX 08-DEC-1999; 99US-0169629P.
PR 06-MAR-2000; 2000US-0187470P.
XX
XX (GEST) GENSET.
XX
XX Dumas Milne Edwards J, Bougueleret L, Jobert S;
PI WPI; 2001-367870/38.
XX P-PSDB; AAC89313.
DR
XX Full length GENSET human nucleic acids encoding potentially secreted
PT proteins, useful in gene therapy and vaccination against a variety of
PT diseases, and for diagnosis of those diseases.
XX
XX Claim 7; Page 753; 921pp; English.
XX
XX The invention relates to full length GENSET human nucleic acids encoding
CC potentially secreted proteins. The nucleic acids and the polypeptides
CC they encode may be used in the prevention, treatment and diagnosis of
CC diseases associated with inappropriate GENSET gene expression. For
CC example, they be used to treat disorders associated with decreased
CC gene expression by rectifying mutations or deletions in a patient's
CC genome that affect the activity of GENSET or by supplementing the
CC patient's own production of GENSET polypeptides. Conversely, antisense
CC nucleic acid molecules may be administered to down regulate GENSET
CC expression by binding with the cells' own genes and preventing their
CC expression. The sense and antisense nucleic acids may also be used as
CC probes in diagnostic assays to detect and quantitate the presence of
CC similar nucleic acid sequences in samples, and hence to determine which
CC patients may be in need of restorative therapy. The GENSET polypeptides
CC may be used as antigens in the production of antibodies and in assays to
CC identify modulators (agonists and antagonists) of GENSET polypeptide
CC expression and activity. The present sequence is a GENSET nucleic acid of
CC the invention
XX
XX Sequence 668 BP; 188 A; 163 C; 151 G; 166 T; 0 U; 0 Other;
SQ
Query Match 5.1%; Score 56; DB 5; Length 668;
Best Local Similarity 100.0%; Pred. No. 2.7e-10;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1034 ATAGCAA 1089
DB 594 ATAGCAA 649
RESULT 8
ABV58410
ID ABV58410 standard; cDNA; 376 BP.
XX
XX ABV58410;
AC
XX
XX 13-SEP-2002 (first entry)
DT
XX Human prostate expression marker cDNA 58401.
DE
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW

XX Pharmacogenomic marker; gene; ss.
XX
XX Homo sapiens.
XX
XX WO200160860-A2.
XX
XX 23-AUG-2001.
XX
XX 20-FEB-2001; 2001WO-US005171.
XX
XX 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX Schlegel R, Endege WO, Monahan JB;
PI WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 11212; 11750pp; English.
XX
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 376 BP; 164 A; 68 C; 61 G; 83 T; 0 U; 0 Other;
SQ
Query Match 5.1%; Score 55; DB 5; Length 376;
Best Local Similarity 100.0%; Pred. No. 6.8e-10;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1035 TAGCAA 1089
DB 215 TAGCAA 269
RESULT 9
AAD02077
ID AAD02077 standard; cDNA; 396 BP.
XX
XX AAD02077;
AC
XX
XX 26-MAR-2001 (first entry)
DT
XX
XX cDNA encoding rat pituitary hormone, pituitrone.
DE
XX Rat; pituitary; pituitrone; therapy; immune disorder; anaemia;
KW Digecore syndrome; haematopoietic cell; ataxia telangiectasia;
KW Wiskott-Aldrich disorder; blood coagulation disorder; Addison's disease;
KW autoimmune disorder; multiple sclerosis; systemic lupus erythematosus;
KW SLE; hyperproliferative disorder; gene therapy; neoplasm;
KW infectious disease; immunomodulatory; cytostatic; antimicrobial; ss.
XX
XX Rattus sp.
OS
XX
XX Key Location/Qualifiers
FH

```

FT CDS 1. .210
FT /*tag= a
FT /product= "Rat pituitary hormone, pituitrone"
FT /note= "Does not include start codon"
FT /partial
XX
XX WO200066778-A1.
XX
XX 09-NOV-2000.
XX
XX 27-APR-2000; 2000WO-US011211.
XX
XX 30-APR-1999; 99US-0131966P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Ruben SM, Ni J;
XX
XX WPI; 2000-687547/67.
XX
XX P-PSDB; AAY1961.
XX
XX Nucleic acids encoding a novel pituitary hormone (pituitrone), useful for
XX diagnosing, preventing and treating e.g. immune disorders,
XX hyperproliferative disorders and blood coagulation disorders.
XX
XX Claim 1; Page 269; 277pp; English.
XX
XX The present sequence is a cDNA encoding rat pituitary hormone,
XX pituitrone. Pituitrone is highly expressed in pituitary gland and is also
XX expressed in brain tissues, spinal cord and kidney. It may be used as
XX antigens in the production of antibodies against pituitrone and in assays
XX to identify modulators. Pituitrone cDNAs are also useful in gene therapy.
XX Pituitrone may be used in the prevention, treatment and diagnosis of
XX diseases associated with inappropriate pituitrone expression. It may be
XX useful in treating disorders related to reproductive and renal system,
XX immune disorders, disorders of haematopoietic cells (e.g., anaemia,
XX George syndrome, ataxia telangiectasia and Wiskott-Aldrich disorder),
XX blood coagulation disorders, autoimmune disorders (e.g. Addison's
XX disease, multiple sclerosis and systemic lupus erythematosus (SLE)),
XX hyperproliferative disorders (e.g. neoplasms of bone, liver and pancreas)
XX and infectious diseases
XX
XX Sequence 396 BP; 110 A; 135 C; 91 G; 60 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 55; DB 3; Length 396;
XX Best Local Similarity 100.0%; Pred. No. 6.7e-10;
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 342 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 396
XX
XX RESULT 10
XX ABX42327/c
XX ID ABX42327 standard; cDNA; 410 BP.
XX
XX AC ABX42327;
XX
XX XX 20-FEB-2003 (first entry)
XX
XX DE Bovine EST associated with lactation/muscle/fat deposition #7492.
XX
XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMPD;
XX muscle deposition; fat deposition; genome mapping; gene identification;
XX gene analysis; cattle breeding.
XX
XX OS Bos Taurus.
XX
XX XX US2002137139-A1.
XX
XX PD 26-SEP-2002.
XX
XX

```

```

PF 24-SEP-2001; 2001US-00960352.
XX
XX 12-JAN-1999; 99US-0115707P.
XX
XX 11-JAN-2000; 2000US-00480902.
XX
XX (BYAT/) BYATT J C.
XX (MATH/) MATHIALAGAN N.
XX (TAON/) TAO N.
XX (WARR/) WARREN W C.
XX
XX Byatt JC, Mathialagan N, Tao N, Warren WC;
XX WPI; 2003-110599/10.
XX
XX New nucleic acid associated with lactation, and muscle and fat
XX deposition, useful for genome mapping, gene identification and analysis,
XX cattle breeding, or for genetically improving cattle.
XX
XX Claim 2; SEQ ID NO 7492; 245pp; English.
XX
XX The invention relates to a purified nucleic acid molecule associated with
XX lactation or muscle and fat deposition (designated LMPD), derived from
XX cattle, and the LMPD nucleic acid can specifically hybridise to a second
XX nucleic acid molecule comprising any of 15112 nucleotide sequences,
XX appearing as ABX34836-ABX49947, or complements of them. Also included are
XX (1) a transformed cell having a nucleic acid comprising an LMPD nucleic
XX acid linked to a promoter and a 3' non-translated sequence that
XX functions in the cell to cause termination of transcription and addition
XX of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
XX (2) determining a level or pattern of a molecule in a bovine cell or
XX tissue comprising: (a) incubating a marker nucleic acid (comprising any
XX of the 15112 nucleic acid sequences or its complement or fragment) with a
XX complementary nucleic acid molecule obtained from the bovine cell or
XX tissue, where hybridisation between the marker nucleic acid and the
XX complementary nucleic acid permits the detection of the molecule; and (b)
XX detecting the level or pattern of the complementary nucleic acid, where
XX the detection of the complementary nucleic acid is predictive of the
XX level or pattern of the molecule. The LMPD nucleic acid is used for
XX determining a level or pattern of a molecule in a bovine cell or tissue.
XX It is useful for genome mapping, gene identification and analysis, cattle
XX breeding, preparation of constructs for use in cattle gene expression, or
XX for genetically improving cattle. The present sequence is one of the
XX 15112 bovine LMPD EST (expressed sequence tag) nucleic acids. Note: The
XX present sequence was not shown in the specification but was obtained in
XX electronic format from the USPTO web site:
XX seqdata.uspto.gov/sequence.html?DocID=20020137139
XX
XX Sequence 410 BP; 151 A; 73 C; 60 G; 126 T; 0 U; 0 Other;
XX
XX Query Match 5.1%; Score 55; DB 8; Length 410;
XX Best Local Similarity 100.0%; Pred. No. 6.6e-10;
XX Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
XX |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
XX Db 80 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 26
XX
XX RESULT 11
XX AA190926
XX ID AA190926 standard; cDNA; 443 BP.
XX
XX AC AA190926;
XX
XX XX 06-NOV-2001 (first entry)
XX
XX DE Human polynucleotide SEQ ID NO 10986.
XX
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX tissue growth factor; immunomodulatory; cancer; leukaemia;
XX nervous system disorders; arthritis; inflammation; ss.
XX

```

OS Homo sapiens.
 PN WO200164835-A2.
 XX
 XX 07-SEP-2001.
 XX
 XX 26-FEB-2001; 2001WO-US004927.
 XX
 XX 28-FEB-2000; 2000US-00515126.
 PR 18-MAY-2000; 2000US-00577409.
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Liu C, Drmanac RT;
 PI WPI; 2001-514838/56;
 DR P-PSDB; AAO10995.
 XX
 XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 PT and treating e.g. leukemia, inflammation and immune disorders.
 PT
 XX Claim 1; SEQ ID NO 10986; 1399pp + Sequence Listing; English.
 XX
 XX The invention relates to human polynucleotides (AA179941-AA193841) and
 CC the encoded proteins (AA00010-AA013910) that exhibit activity elating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or
 CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
 CC inflammation. Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic format
 CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 443 BP; 183 A; 68 C; 96 G; 96 T; 0 U; 0 Other;
 Query Match 5.1%; Score 55; DB 4; Length 443;
 Best Local Similarity 100.0%; Pred. No. 6.5e-10;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1035 TAGCAA 1089
 DB 137 TAGCAA 191
 RESULT 12
 AAZ52570
 ID AAZ52570 standard; cDNA; 491 BP.
 XX
 XX AAZ52570;
 AC
 XX 29-FEB-2000 (first entry)
 DT
 XX Human secreted protein clone yd261_1 nucleotide sequence SEQ ID NO:191.
 DE
 XX Human; secreted protein; immunostimulatory; haemostatic; cytokine;
 KW proliferative; differentiative; chemotactic; chemokinetic; vaccine;
 KW thrombolytic; antiinflammatory; cytostatic; immunosuppressive;
 KW gene therapy; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9958642-A2.
 XX
 XX 18-NOV-1999.
 PD
 XX 14-MAY-1999; 99WO-US010843.
 PF
 XX 14-MAY-1998; 98US-0085472P.
 PR 17-AUG-1998; 98US-0096824P.
 PR

PR 11-SEP-1998; 98US-0099843P.
 PR 11-SEP-1998; 98US-0099950P.
 PR 15-SEP-1998; 98US-0100424P.
 PR 29-SEP-1998; 98US-0102329P.
 PR 09-OCT-1998; 98US-0103615P.
 PR 11-DEC-1998; 98US-0111779P.
 PR 14-DEC-1998; 98US-0112159P.
 PR 31-DEC-1998; 98US-0114415P.
 PR 10-FEB-1999; 99US-00248059.
 PR 06-APR-1999; 99US-00287150.
 PR 13-MAY-1999; 99US-00311021.
 XX (GEMY) GENETICS INST INC.
 PA
 XX Wong GG, Clark HF, Fechtel K, Agostino MJ;
 PI WPI; 2000-053095/04.
 DR P-PSDB; AAY73485.
 XX
 XX Novel polynucleotides and proteins having biological activities which
 PT make them suitable for treating, preventing or ameliorating medical
 PT conditions in humans or animals.
 PT
 XX Claim 200; Page 702; 730pp; English.
 PS
 XX The present invention describes human secreted proteins encoded by
 CC polynucleotides obtained from adult testes, foetal brain, adult brain,
 CC brain (foetal and adult), foetal kidney, adult spleen, and adult thymus
 CC cDNA libraries. The polynucleotides and proteins are predicted to have
 CC biological activities which would make them suitable for treating,
 CC preventing or ameliorating medical conditions in humans and animals.
 CC Suggested activities include nutritional activity, cytokine and cell
 CC proliferation/differentiation activity, immune stimulating (e.g. as
 CC vaccines) or suppressing activity, haematopoiesis regulating activity,
 CC tissue growth activity, activin/inhibin activity, chemotactic/
 CC chemokinetic activity, haemostatic and thrombolytic activity, receptor/
 CC ligand activity, anti-inflammatory activity, cadherin/tumour invasion
 CC suppressor activity, and tumour inhibition activity. The polynucleotides
 CC are also stated to be useful for gene therapy. Therapeutic compositions
 CC are also presently valuable for veterinary applications. AAZ52475 to
 CC AAZ52581 encode human secreted proteins, and AAY73390 to AAY73500
 CC represent human secreted proteins, given in the present invention
 XX
 SQ Sequence 491 BP; 133 A; 148 C; 114 G; 96 T; 0 U; 0 Other;
 Query Match 5.1%; Score 55; DB 3; Length 491;
 Best Local Similarity 100.0%; Pred. No. 6.4e-10;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 1035 TAGCAA 1089
 DB 430 TAGCAA 484
 RESULT 13
 AAC77635
 ID AAC77635 standard; cDNA; 917 BP.
 XX
 XX AAC77635;
 AC
 XX 08-FEB-2001 (first entry)
 DT
 XX Human cancer associated gene sequence SEQ ID NO:29.
 DE
 XX Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; cancer antigen; immunomodulator;
 KW antiidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiac;
 KW dermatological; neuroprotective; thrombolytic; coagulant; nootropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;

```

KW neurological disease; drug screening; ss.
XX
OS Homo sapiens.
XX
PN WO200055350-A1.
XX
PD 21-SEP-2000.
XX
PF 08-MAR-2000; 2000WO-US005882.
XX
PR 12-MAR-1999; 99US-0124270P.
XX
PA (HUMA-) HUMAN GENOME SCI INC.
XX
PI Rosen CA, Ruben SM;
DR WPI; 2000-587533/55.
DR P-PSDB; AAB43426.
XX
XX Novel isolated nucleic acids comprising sequences encoding peptides
PT useful for treating or diagnosing e.g. cancer.
PT
PS Claim 1; Page 632; 2352pp; English.
XX
XX AAC77607 to AAC78448 encode the human cancer associated proteins given in
CC AAB43398 to AAB44239. The proteins can have activities based on the
CC tissues and cells the genes are expressed in. Example of activities
CC include: cytostatic; proliferative; tumorigenic; immunomodulatory;
CC antidiabetic; antiasthmatic; antirheumatic; antiarthritic;
CC antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;
CC dermatological; neuroprotective; cardiant; thrombolytic; coagulant;
CC neotropic; vasotropic; antipsoriatic and antiangiogenic. The
CC polynucleotides and polypeptides can be used for preventing, treating or
CC ameliorating medical conditions and diagnosing pathological conditions.
CC Polynucleotides, polypeptides, antibodies, agonists and antagonists from
CC the present invention may be used to treat immune disorders by activating
CC or inhibiting the proliferation, differentiation or mobilisation of
CC immune cells, to treat disorders of haematopoietic cells, autoimmune
CC rejection, allergic reactions, graft versus host disease and organ
CC disorders, modulate haemostatic or thrombolytic activity, modulate
CC inflammation, cancers, cardiovascular disorders, neurological disease and
CC bacterial or viral infections. The peptides, nucleotides, antibodies,
CC agonists and antagonists may be also be used in drug screens. AAC78449 to
CC AAC78457 and AAB44240 represent sequences used in the exemplification of
CC the present invention
XX
SQ Sequence 917 BP; 247 A; 269 C; 257 G; 142 T; 0 U; 2 Other;
Query Match 5.1%; Score 55; DB 3; Length 917;
Best Local Similarity 100.0%; Pred. No. 5.5e-10;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 841 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 895
RESULT 14
ADQ22529
ID ADQ22529 standard; DNA; 2435 BP.
XX
XX ADQ22529;
AC ADQ22529;
XX
XX 26-AUG-2004 (first entry)
DT
DE Human soft tissue sarcoma-upregulated DNA - SEQ ID 5349.
XX
XX soft tissue sarcoma; cytostatic; gene therapy; vaccine; screening; human;
KW db.
XX
XX Homo sapiens.
OS
XX
XX WO2004048938-A2.
PN

```

```

XX 10-JUN-2004.
PD
XX
XX 26-NOV-2003; 2003WO-US038193.
XX
XX 26-NOV-2002; 2002US-0429739P.
XX
XX (PROT-) PROTEIN DESIGN LABS INC.
PA
XX
XX Aziz N, Ginsburg WM, Zlotnik A;
PI WPI; 2004-441208/41.
DR
XX
XX Early detection of soft tissue sarcoma comprises determining expression
PT of a gene in a first soft tissue sample and a normal soft tissue sample
PT and comparing the gene expression, also useful in treating soft tissue
PT sarcoma.
XX
XX Example 2; SEQ ID NO 5349; 210pp; English.
PS
XX
XX The invention relates to a novel method for detecting soft tissue sarcoma
CC which comprises obtaining a first soft tissue sample from an individual
CC and a normal soft tissue sample from the same or different individual,
CC determining the expression of a gene in both samples and comparing the
CC expression of the gene in both soft tissue samples, where a higher level
CC of protein expression in the first soft tissue sample indicates the
CC presence of soft tissue sarcoma. The method of the invention has
CC cytostatic applications and may be useful for detecting soft tissue
CC sarcoma, possibly via gene therapy or vaccine production. The nucleic
CC acid sequences may be useful in diagnostic and screening applications.
CC The current sequence is that of a human soft tissue sarcoma-upregulated
CC DNA of the invention. The current sequence is not shown within the
CC specification per se but was submitted in CD format by the inventor.
XX
SQ Sequence 2435 BP; 637 A; 571 C; 538 G; 688 T; 0 U; 1 Other;
Query Match 5.1%; Score 55; DB 12; Length 2435;
Best Local Similarity 100.0%; Pred. No. 4.4e-10;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
DB 2358 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2412
RESULT 15
AAC58600
ID AAC58600 standard; cDNA; 3759 BP.
XX
XX AAC58600;
AC AAC58600;
XX
XX 29-JAN-2001 (first entry)
DT
DE Human PRO1083 protein UNQ540 encoding cDNA SEQ ID NO:116.
XX
XX Human; immune related disease; diagnosis; antiinflammatory; cardiant;
KW dermatological; antiarthritic; antirheumatic; immunosuppressive;
KW haemostatic; antithyroid; antidiabetic; neotropic; neuroprotective;
KW antianaemic; hepatotropic; viricide; antipsoriatic; antiallergic;
KW antiasthmatic; systemic lupus erythematosus; rheumatoid arthritis;
KW osteoarthritis; spondyloarthropathy; systemic sclerosis; sarcoidosis;
KW idiopathic inflammatory myopathy; Sjogren's syndrome; thyroiditis;
KW systemic vasculitis; autoimmune haemolytic anaemia; diabetes mellitus;
KW autoimmune thrombocytopaenia; immune-mediated renal disease;
KW demyelinating disease; hepatobiliary disease; Whipple's disease;
KW inflammatory bowel disease; gluten-sensitive enteropathy;
KW autoimmune disease; immune-mediated skin disease; allergic disease;
KW immunological disease; transplantation associated disease;
KW graft rejection; graft-versus-host-disease; ss.
XX
XX Homo sapiens.
OS
XX
XX WO200053758-A2.
PN

```

XX 14-SEP-2000.
 XX 02-MAR-2000; 2000WO-US005841.
 XX 08-MAR-1999; 99WO-US005028.
 XX 10-MAR-1999; 99US-0123618P.
 XX 12-MAR-1999; 99US-0123957P.
 XX 23-MAR-1999; 99US-0125775P.
 XX 12-APR-1999; 99US-0128849P.
 XX 20-APR-1999; 99WO-US008615.
 XX 28-APR-1999; 99US-0131445P.
 XX 04-MAY-1999; 99US-0132371P.
 XX 14-MAY-1999; 99US-0134287P.
 XX 02-JUN-1999; 99WO-US012252.
 XX 23-JUN-1999; 99US-0141037P.
 XX 20-JUL-1999; 99US-0144758P.
 XX 26-JUL-1999; 99US-0145698P.
 XX 28-JUL-1999; 99US-0146222P.
 XX 01-SEP-1999; 99WO-US020111.
 XX 08-SEP-1999; 99WO-US020594.
 XX 13-SEP-1999; 99WO-US020944.
 XX 15-SEP-1999; 99WO-US021090.
 XX 15-SEP-1999; 99WO-US021547.
 XX 05-OCT-1999; 99WO-US023089.
 XX 29-OCT-1999; 99US-0162506P.
 XX 29-NOV-1999; 99WO-US028214.
 XX 30-NOV-1999; 99WO-US028313.
 XX 30-NOV-1999; 99WO-US028409.
 XX 01-DEC-1999; 99WO-US028301.
 XX 01-DEC-1999; 99WO-US028634.
 XX 02-DEC-1999; 99WO-US028551.
 XX 02-DEC-1999; 99WO-US028564.
 XX 02-DEC-1999; 99WO-US028565.
 XX 16-DEC-1999; 99WO-US030095.
 XX 20-DEC-1999; 99WO-US030999.
 XX 20-DEC-1999; 99WO-US031274.
 XX 05-JAN-2000; 2000WO-US000219.
 XX 06-JAN-2000; 2000WO-US000277.
 XX 06-JAN-2000; 2000WO-US000376.
 XX 11-FEB-2000; 2000WO-US003565.
 XX 18-FEB-2000; 2000WO-US004341.
 XX 18-FEB-2000; 2000WO-US004342.
 XX 22-FEB-2000; 2000WO-US004414.

(GETH) GENENTECH INC.

XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hebert C, Henzel W;
 PI Kabakoff RC, Lu Y, Pan J, Pennica D, Shelton DL, Smith V;
 PI Stewart TA, Tumas D, Watanabe CK, Wood WI, Yan M;

DR WPI: 2000-572271/53.
 DR P-PSDB; AAB33435.

XX Sixty four PRO polypeptides, useful in the diagnosis and treatment of
 PT immune related disorders, e.g. systemic lupus erythematosus, rheumatoid
 PT arthritis, osteoarthritis, thyroiditis and diabetes mellitus.

XX Claim 23; Fig 43; 309pp; English.

XX The present invention describes sixty four human PRO proteins which can
 CC be used in the treatment of immune related diseases. The human PRO
 CC proteins, anti-PRO antibodies, agonists and antagonists are useful for
 CC treating and diagnosing immune related disorders. The disorders are
 CC selected from systemic lupus erythematosus, rheumatoid arthritis,
 CC osteoarthritis, juvenile chronic arthritis, spondyloarthritis,
 CC systemic sclerosis, idiopathic inflammatory myopathies, Sjogren's
 CC syndrome, systemic vasculitis, sarcoidosis, autoimmune haemolytic
 CC anaemia, autoimmune thrombocytopenia, thyroiditis, diabetes mellitus,
 CC immune-mediated renal disease, demyelinating diseases of the central and
 CC peripheral nervous systems, hepatobiliary diseases, inflammatory bowel
 CC disease, gluten-sensitive enteropathy and Whipple's disease, autoimmune
 CC or immune-mediated skin diseases, allergic diseases, immunological

CC diseases of the lung, and transplantation associated diseases including
 CC graft rejection and graft-versus-host-disease. AAC58397 to AAC58578
 CC represent PCR primers and hybridisation probes used in the isolation of
 CC human PRO sequences. AAC58579 to AAC58642 and AAB33414 to AAB33477
 CC represent human PRO polynucleotide and protein sequences given in the
 CC exemplification of the present invention
 XX
 SQ Sequence 3759 BP; 774 A; 1205 C; 970 G; 810 T; 0 U; 0 Other;

Query Match 5.1%; Score 55; DB 3; Length 3759;
 Best Local Similarity 100.0%; Pred. No. 4e-10;
 Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089
 ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
 Db 3675 TAGCAAA 3729

Search completed: February 6, 2005, 17:35:48
 Job time : 704 secs

This Page Blank (usp10)

1	59	5.4	1199	18	US-10-425-115-36465	Sequence 36465, A
2	56	5.1	668	9	US-09-731-8972-192	Sequence 192, App
3	56	5.1	668	10	US-09-876-997-192	Sequence 192, App
4	56	5.1	2121	18	US-10-425-115-184240	Sequence 184240, App
5	55	5.1	222	18	US-10-425-115-150417	Sequence 150417, A
6	55	5.1	369	18	US-10-425-115-52629	Sequence 52629, A
7	55	5.1	376	9	US-10-357-930-58429	Sequence 58429, A
8	55	5.1	410	9	US-09-960-352-7492	Sequence 7492, Ap
9	55	5.1	538	18	US-10-425-115-91860	Sequence 91860, A
10	55	5.1	787	17	US-10-424-599-122506	Sequence 122506, A
11	55	5.1	838	18	US-10-425-115-163680	Sequence 163680, A

US-09-731-872-192

Pred. No.	is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.	CURR	NUMBER	SEQ ID	LENGTH

; Sequence 192, Application US/09731872
; Patent No. US20020102604A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US3 REG
; CURRENT APPLICATION NUMBER: US/09/731,872
; CURRENT FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/169,629
; PRIOR FILING DATE: 1999-12-08
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 192
; LENGTH: 668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..203
US-09-731-872-192

Query Match 5.1%; Score 56; DB 9; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089
DB 594 ATAGCAA 649

RESULT 3
US-09-876-997-192
; Sequence 192, Application US/09876997
; Publication No. US20030152921A1
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, Jean Baptiste
; APPLICANT: Bougueleret, Lydie
; APPLICANT: Jobert, Severin
; TITLE OF INVENTION: FULL-LENGTH HUMAN CDNAS ENCODING POTENTIALLY SECRETED PROTEINS
; FILE REFERENCE: 78 US4 CIP
; CURRENT APPLICATION NUMBER: US/09/876,997
; CURRENT FILING DATE: 2001-06-08
; PRIOR APPLICATION NUMBER: US 09/731,872
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: US 60/187,470
; PRIOR FILING DATE: 2000-03-06
; PRIOR APPLICATION NUMBER: US 60/169,629
; NUMBER OF SEQ ID NOS: 482
; SOFTWARE: Patent.pm
; SEQ ID NO 192
; LENGTH: 668
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 57..203
US-09-876-997-192

Query Match 5.1%; Score 56; DB 10; Length 668;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089
DB 594 ATAGCAA 649

RESULT 4

US-10-425-115-184240
; Sequence 184240, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 184240
; LENGTH: 2121
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_99614C.1
US-10-425-115-184240

Query Match 5.1%; Score 56; DB 18; Length 2121;
Best Local Similarity 100.0%; Pred. No. 1.6e-17;
Matches 56; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1034 ATAGCAA 1089
DB 1902 ATAGCAA 1957

RESULT 5
US-10-425-115-150417/c
; Sequence 150417, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 150417
; LENGTH: 222
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_68706C.1
US-10-425-115-150417

Query Match 5.1%; Score 55; DB 18; Length 222;
Best Local Similarity 100.0%; Pred. No. 5.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAA 1089
DB 75 TAGCAA 21

RESULT 6
US-10-425-115-52629
; Sequence 52629, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants
/ FILE REFERENCE: 38-21(53222)B
/ CURRENT APPLICATION NUMBER: US/10/425,115
/ CURRENT FILING DATE: 2003-04-28
/ NUMBER OF SEQ ID NOS: 369326
/ SEQ ID NO 52629
/ LENGTH: 369
/ TYPE: DNA
/ ORGANISM: Zea mays
/ FEATURE:
/ OTHER INFORMATION: Clone ID: MRT4577_147996C.1
US-10-425-115-52629

Query Match 5.1%; Score 55; DB 18; Length 369;
Best Local Similarity 100.0%; Pred. No. 5.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089
|||
DB 129 TAGCAAA 183

RESULT 7

US-10-357-930-58429
/ Sequence 58429, Application US/10357930
/ Publication No. US20040259086A1
/ GENERAL INFORMATION:

/ APPLICANT: Schlegel, Robert
/ APPLICANT: Monahan, John
/ APPLICANT: Schlegel, Robert

/ TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
/ TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
/ TITLE OF INVENTION: HUMAN PROSTATE CANCER

/ FILE REFERENCE: MRI-007BCN
/ CURRENT APPLICATION NUMBER: US/10/357,930
/ CURRENT FILING DATE: 2003-02-04

/ PRIOR APPLICATION NUMBER: 09/785,276
/ PRIOR FILING DATE: 2003-02-16

/ PRIOR APPLICATION NUMBER: 60/183,319
/ PRIOR FILING DATE: 2000-02-17

/ PRIOR APPLICATION NUMBER: 60/189,862
/ PRIOR FILING DATE: 2000-03-16

/ PRIOR APPLICATION NUMBER: 60/207,454
/ PRIOR FILING DATE: 2000-05-25

/ PRIOR APPLICATION NUMBER: 60/211,314
/ PRIOR FILING DATE: 2000-06-09

/ PRIOR APPLICATION NUMBER: 60/219,007
/ PRIOR FILING DATE: 2000-07-18

/ PRIOR APPLICATION NUMBER: 60/255,281
/ PRIOR FILING DATE: 2000-12-13

/ NUMBER OF SEQ ID NOS: 62232
/ SOFTWARE: FastSeq for Windows Version 4.0

/ SEQ ID NO 58429
/ LENGTH: 376
/ TYPE: DNA

/ ORGANISM: Homo sapiens

US-10-357-930-58429

Query Match 5.1%; Score 55; DB 18; Length 376;
Best Local Similarity 100.0%; Pred. No. 5.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089
|||
DB 215 TAGCAAA 269

RESULT 8

US-09-960-352-7492/C
/ Sequence 7492, Application US/09960352
/ Patent No. US20020137139A1
/ GENERAL INFORMATION:

/ APPLICANT: Warren, Wesley C.

/ APPLICANT: Tao, Nengbing
/ APPLICANT: Byatt, John C.
/ APPLICANT: Mathialagan, Nagappan
/ TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
/ TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
/ FILE REFERENCE: 16511.006/37-21(10298)C
/ CURRENT APPLICATION NUMBER: US/09/960,352
/ CURRENT FILING DATE: 2001-09-24
/ NUMBER OF SEQ ID NOS: 15112
/ SEQ ID NO 7492
/ LENGTH: 410
/ TYPE: DNA
/ ORGANISM: Bos taurus
/ OTHER INFORMATION: Clone ID: 32-LIB34-017-Q1-E1-H7
US-09-960-352-7492

Query Match 5.1%; Score 55; DB 9; Length 410;
Best Local Similarity 100.0%; Pred. No. 5.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089
|||
DB 80 TAGCAAA 26

RESULT 9

US-10-425-115-91860/C

/ Sequence 91860, Application US/10425115

/ Publication No. US20040214272A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J.

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants

/ FILE REFERENCE: 38-21(53222)B

/ CURRENT APPLICATION NUMBER: US/10/425,115

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 369326

/ SEQ ID NO 91860

/ LENGTH: 538

/ TYPE: DNA

/ ORGANISM: Zea mays

/ FEATURE:

/ OTHER INFORMATION: Clone ID: MRT4577_183771C.1

US-10-425-115-91860

Query Match 5.1%; Score 55; DB 18; Length 538;
Best Local Similarity 100.0%; Pred. No. 5.1e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAA 1089
|||
DB 96 TAGCAAA 42

RESULT 10

US-10-424-599-122506

/ Sequence 122506, Application US/10424599

/ Publication No. US20040031072A1

/ GENERAL INFORMATION:

/ APPLICANT: La Rosa, Thomas J.

/ APPLICANT: Kovalic, David K.

/ APPLICANT: Zhou, Yihua

/ APPLICANT: Cao, Yongwei

/ TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

/ TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement

/ FILE REFERENCE: 38-21(53223)B

/ CURRENT APPLICATION NUMBER: US/10/424,599

/ CURRENT FILING DATE: 2003-04-28

/ NUMBER OF SEQ ID NOS: 285684

/ SEQ ID NO 122506

```
; LENGTH: 787
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_81629C.1
US-10-424-599-122506

Query Match          5.1%; Score 55; DB 17; Length 787;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 672 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 726

RESULT 11
US-10-425-115-163680
; Sequence 163680, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 163680
; LENGTH: 838
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(838)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_80852C.1
US-10-425-115-163680

Query Match          5.1%; Score 55; DB 18; Length 838;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 670 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 724

RESULT 12
US-09-925-301-29
; Sequence 29, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 29
; LENGTH: 917
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-29

Query Match          5.1%; Score 55; DB 9; Length 917;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 841 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 895

RESULT 13
US-09-803-589-13
; Sequence 13, Application US/09803589
; Patent No. US20020112251A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-325001
; CURRENT APPLICATION NUMBER: US/09/803,589
; CURRENT FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(630)
US-09-803-589-13

Query Match          5.1%; Score 55; DB 9; Length 1027;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1035 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1089
Db 958 TAGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1012

RESULT 14
US-10-718-332-13
; Sequence 13, Application US/10718332
; Publication No. US20040253605A1
; GENERAL INFORMATION:
; APPLICANT: McCarthy, Sean A.
; APPLICANT: Holtzman, Douglas A.
; APPLICANT: Goodearl, Andrew D.J.
; TITLE OF INVENTION: NOVEL GENES ENCODING PROTEINS HAVING
; TITLE OF INVENTION: PROGNOSTIC, DIAGNOSTIC, PREVENTIVE, THERAPEUTIC AND OTHER
; TITLE OF INVENTION: USES
; FILE REFERENCE: 07334-325001
; CURRENT APPLICATION NUMBER: US/10/718,332
; CURRENT FILING DATE: 2003-11-20
; PRIOR APPLICATION NUMBER: US/09/803,589
```

;
; PRIOR FILING DATE: 2001-03-09
; PRIOR APPLICATION NUMBER: US 09/128,709
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: US 60/054,645
; PRIOR FILING DATE: 1997-08-04
; PRIOR APPLICATION NUMBER: US 09/130,491
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: US 60/054,966
; PRIOR FILING DATE: 1997-08-06
; PRIOR APPLICATION NUMBER: US 60/058,108
; PRIOR FILING DATE: 1997-09-05
; PRIOR APPLICATION NUMBER: US 09/388,280
; PRIOR FILING DATE: 1999-09-01
; PRIOR APPLICATION NUMBER: US 09/388,279
; PRIOR FILING DATE: 1999-09-01
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 13
; LENGTH: 1027
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (106)...(630)
US-10-718-332-13

Query Match 5.1%; Score 55; DB 18; Length 1027;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1035 TAGCAAA 1089
DB 958 TAGCAAA 1012

RESULT 15
US-10-723-860-5349
; Sequence 5349, Application US/10723860
; Publication No. US20040253606A1
; GENERAL INFORMATION:
; APPLICANT: Aziz, Natasha
; APPLICANT: Ginsburg, Wendy M.
; APPLICANT: Zlotnik, Albert
; TITLE OF INVENTION: Methods of Diagnosis of Soft Tissue Sarcoma, Compositions &
; FILE REFERENCE: 05892.0193.NPUS01
; CURRENT APPLICATION NUMBER: US/10/723,860
; CURRENT FILING DATE: 2003-11-26
; PRIOR APPLICATION NUMBER: 60/429,739
; PRIOR FILING DATE: 2002-11-26
; NUMBER OF SEQ ID NOS: 8393
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 5349
; LENGTH: 2435
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (2063)..(2063)
; OTHER INFORMATION: n is a, c, g, or t
US-10-723-860-5349

Query Match 5.1%; Score 55; DB 18; Length 2435;
Best Local Similarity 100.0%; Pred. No. 5.2e-17;
Matches 55; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1035 TAGCAAA 1089
DB 2358 TAGCAAA 2412

Search completed: February 6, 2005, 21:54:29
Job time : 692 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:47:21 ; Search time 5146 Seconds
(without alignments)
10254.126 Million cell updates/sec

Title: US-10-088-830-1
Perfect score: 1089
Sequence: 1 gaattcgacgagccgca.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*
2: gb_hg.*
3: gb_in.*
4: gb_cm.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1084.2	99.6	1089	6	AX100704 Sequence
2	1065.4	97.8	1083	8	TSP271917
3	515	47.3	1174	8	AK111611 Oryza sat
4	442.2	40.6	885	8	AY224529
5	168	15.4	168	6	AX100712 Sequence
6	168	15.4	168	6	AX100714 Sequence
7	153.6	14.1	879	8	BT005286 Arabidops
8	153.6	14.1	960	8	ATH319027 Arabidops
9	153.6	14.1	1114	6	AX449297 Sequence
10	153.6	14.1	1121	8	AK117135 Arabidops
11	153.6	14.1	1189	6	AX449329 Sequence
12	153.6	14.1	1274	6	AX712207 Sequence
13	153.6	14.1	1274	6	AX734251 Sequence
14	153.6	14.1	1274	8	ATH294531 Arabidops
15	153	13.9	153	6	AX100710 Sequence
16	151.2	13.9	156	6	AX100708 Sequence
17	150.8	13.8	1158	8	ATH294532 Arabidops
18	150.8	13.8	1618	8	AY086018 Arabidops
19	150.6	13.8	1640	8	AF181998 Populus t

20	150.6	13.8	1640	8	AY307373
21	150	13.8	150	6	AX100716 Sequence
22	136.2	12.5	1140	8	AY224589 Oryza sat
23	132.2	12.1	1041	8	AY224551 Oryza sat
24	120.6	11.1	19459	8	AP003335 Oryza sat
25	118.6	10.9	1131	6	AX449314 Sequence
26	104.6	9.6	426	6	AX449520 Sequence
27	98.8	9.1	1313	9	BC021113 Homo sapi
28	98.8	9.1	1802	9	BC013993 Homo sapi
29	98.8	9.1	2277	6	CQ17592 Sequence
30	98.8	9.1	2320	6	AR380911 Sequence
31	98.8	9.1	2320	9	HSDE2 Homo sapien
32	97.6	9.0	1969	5	BC082841 Xenopus 1
33	97	8.9	1266	6	AX333694 Sequence
34	97	8.9	1266	9	HUMDP2M Human DP-2
35	97	8.9	2968	6	CQ489116 Sequence
36	97	8.9	2968	6	CQ494960 Sequence
37	94.4	8.7	1902	5	BC080383 Xenopus 1
38	93.6	8.6	870	6	AX449515 Sequence
39	93.6	8.6	1442	6	AX449342 Sequence
40	88	8.1	1226	3	DMDPRTF Drosophila
41	88	8.1	2200	3	AY028165 Casnorhab
42	88	8.1	2203	6	CQ581589 Sequence
43	88	8.1	2461	3	BT003183 Drosophil
44	88	8.1	2495	3	AY069526 Drosophil
45	86.6	8.0	2325	5	BC044409 Danio rer

ALIGNMENTS

RESULT 1	AX100704	1089 bp	DNA	linear	PAT 10-APR-2001
LOCUS	Sequence 1 from Patent WO0121644.				
DEFINITION	AX100704				
ACCESSION	AX100704.1				
VERSION	GI:13619652				
KEYWORDS	Triticum monococcum				
SOURCE	Triticum monococcum				
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Poideae; Triticeae; Triticum.				
REFERENCE	1				
AUTHORS	Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.				
TITLE	Wheat dp proteins and uses thereof				
JOURNAL	Patent: WO 0121644-A 1 29-MAR-2001;				
FEATURES	CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)				
source	Location/Qualifiers				
	1..1089				
	/organism="Triticum monococcum"				
	/mol_type="unassigned DNA"				
	/db_xref="taxon:4568"				
	20..805				
	/notes="unnamed protein product"				
	/codon_start=1				
	/protein_id="CAC36466.1"				
	/db_xref="GI:13619653"				
	/translation="MAPPPGAAAAATAALDTGVHILEASSVPPLPERGNAVQRKG AVDPDKRKKKAAAPRITGWLREYSKIVCEKVEAKGRITTYNEVADEITYSELKMAH IQGFDEKIRRRVDAFNVLIALRVIAKEKEIRWGLSNRYEKIKLEEVKELV NKIRNKALLOIEKQFDLONIKLNQTLSSAENVNGIRLPFLVLTSKARVEIE ISDSKFAHFEENGAPFTLHDDLILGVRNRSIGRATLH"				

ORIGIN

Query Match	99.6%;	Score 1084.2;	DB 6;	Length 1089;
Best Local Similarity	99.7%;	Pred. No. 9.4e-229;		
Matches 1086;	Conservative	0;	Mismatches 3;	Indels 0;
Gaps	0;			
QY	1	GAATTCGGACGAGCGCGCAATGCGCGCTCCCGCGCGGAGCTGCTGCGCGCGCTACCGC	60	
Db	1	GAATTCGGACGAGCGCGCAATGCGCGCTCCCGCGCGGAGCTGCTGCGCGCGCTACCGC	60	

```
QY 61 CGACGTGGACCTGACCGCGGTGCACATCTCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGA 120
Db 61 CGACGTGGACCTGACCGCGGTGCACATCTCTCGAAGCTTCCAGTGTCCCGCGCTTCCCGA 120
QY 121 AGCGGGCGGTAAATCGCGTCCAAAGGAAGGGGGCTGTTGACCCGGATAAAGATAGGAAGAA 180
Db 121 AGCGGGCGGTAAATCGCGTCCAAAGGAAGGGGGCTGTTGACCCGGATAAAGATAGGAAGAA 180
QY 181 GGAAGAGCTCGCGCACCGAGGATCACCGTGTGGGGCTCCGGAGTACAGCAAAATAGT 240
Db 181 GGAAGAGCTCGCGCACCGAGGATCACCGTGTGGGGCTCCGGAGTACAGCAAAATAGT 240
QY 241 TTGTGAGAAAGTTGAAGCCAAAGGAAGCAACAATCAATGAGTGTGCAGACCAAAATTTA 300
Db 241 TTGTGAGAAAGTTGAAGCCAAAGGAAGCAACAATCAATGAGTGTGCAGACCAAAATTTA 300
QY 301 TTGAGAGCTGAAGTCCATGCGCACATATTGCTCAAGGGTGTGATGAGAAGAAATATTAGCG 360
Db 301 TTGAGAGCTGAAGTCCATGCGCACATATTGCTCAAGGGTGTGATGAGAAGAAATATTAGCG 360
QY 361 GAGAGTGTATGATGCTTTCAACGTTCTCATTTGACCTTCGTTGTTATTCGCAAAAGAAAAA 420
Db 361 GAGAGTGTATGATGCTTTCAACGTTCTCATTTGACCTTCGTTGTTATTCGCAAAAGAAAAA 420
QY 421 GGAATACGCTGGATGGCGCTTTCAAAATTAAGATATGAAAAATTAAGAGCTTGAGGA 480
Db 421 GGAATACGCTGGATGGCGCTTTCAAAATTAAGATATGAAAAATTAAGAGCTTGAGGA 480
QY 481 AGTTGCTAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGA 540
Db 481 AGTTGCTAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGGAATTCGA 540
QY 541 AAAACAGTTTGTATGATCTCAAAACATCAAGTTACGTAACAAACACTGGAAGCTCAGC 600
Db 541 AAAACAGTTTGTATGATCTCAAAACATCAAGTTACGTAACAAACACTGGAAGCTCAGC 600
QY 601 AGAGATGTTAATGGCATCGCCTTCCATTCGTTTGGTCAAGACATCTAGGAAGCAAG 660
Db 601 AGAGATGTTAATGGCATCGCCTTCCATTCGTTTGGTCAAGACATCTAGGAAGCAAG 660
QY 661 GGTGGAATTTGAGATTTGAGATGACTGCAAGTTTGGCCATTTCCAGTTTCAATGGTGACC 720
Db 661 GGTGGAATTTGAGATTTGAGATGACTGCAAGTTTGGCCATTTCCAGTTTCAATGGTGACC 720
QY 721 ATTACATTCGATGATCTCTCAATCCTTGGAGGGGTAAAGCGTAAACGATAGGAAG 780
Db 721 ATTACATTCGATGATCTCTCAATCCTTGGAGGGGTAAAGCGTAAACGATAGGAAG 780
QY 781 AGCTGGCGCGCCACCTTCACTAGACATCAAGATATTACAAATGAATTTAAAGTGT 840
Db 781 AGCTGGCGCGCCACCTTCACTAGACATCAAGATATTACAAATGAATTTAAAGTGT 840
QY 841 AGAAGTGGCAGCGCGGATTTCTTTGACAGCTATGATAGCTATATATCTCATGAAAA 900
Db 841 AGAAGTGGCAGCGCGGATTTCTTTGACAGCTATGATAGCTATATATCTCATGAAAA 900
QY 901 CTGACCTAGTTTATAGGACAGTCTCTCAGCTTGAGAGATTTTAACTGCAAAATTTTG 960
Db 901 CTGACCTAGTTTATAGGACAGTCTCTCAGCTTGAGAGATTTTAACTGCAAAATTTTG 960
QY 961 TCTCCTTTTGTGCTAGCAGGTTATTAGTCTCAGATAGATGATTCATATATGCTGCTG 1020
Db 961 TCTCCTTTTGTGCTAGCAGGTTATTAGTCTCAGATAGATGATTCATATATGCTGCTG 1020
QY 1021 TATGAAAAACATTGATAGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1080
Db 1021 TATGAAAAACATTGATAGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1080
QY 1081 AAAAAAAA 1089
Db 1081 AAAAAAAA 1089
```

```
RESULT 2
TSP271917
LOCUS TSP271917 1083 bp mRNA linear PLN 16-DEC-2000
DEFINITION Triticum sp. mRNA for DP protein (dp gene).
ACCESSION AJ271917
VERSION AJ271917.1 GI:11877790
KEYWORDS DP gene; E2F dimerization partner.
SOURCE Triticum sp.
ORGANISM Triticum sp.
REFERENCE 1
AUTHORS Ramirez-Parra, F. and Gutierrez, C.
TITLE Characterization of wheat DP, a heterodimerization partner of the
JOURNAL Plant E2F transcription factor which stimulates E2F-DNA binding
MEDLINE FEBS Lett. 486 (1), 73-78 (2000)
PUBMED 20562804
REFERENCE 2 (bases 1 to 1083)
AUTHORS Gutierrez, C.
TITLE Direct Submission
JOURNAL Submitted (03-FEB-2000) Gutierrez C., C.S.I.C.-U.A.M., Centro de
Biologia Molecular, Cantoblanco, Madrid 28049, SPAIN
FEATURES
source
1..1083
/mol_type="mRNA"
/db_xref="taxon:4569"
1..1083
/gene="dp"
<1..13
/gene="dp"
14..799
/gene="dp"
/function="cell cycle E2F dimerization partner"
/codon_start=1
/product="DP protein"
/protein_id="CAC19034.1"
/db_xref="GI:11877791"
/db_xref="GOA:Q9FET1"
/db_xref="UniProt/TrEMBL:Q9FET1"
/translat="MAPPRGAAAAAALDLTGVLHLEASSVPPLPERGNAVORRG
AVDPDKRKEKAAPRTGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELSMAH
IGGPFDEKIRRVDAFNVLIALRVIAKEKEIRWNGISNYREYKIKLEEVKELV
NKIRKKALLQEIEKQFDLQNKLRNQLSESAENVNGIRLPFLVLTYSRKARVELE
ISDSKFAHFEFNGAPFTLLHDDLILLEGVRGNSIGKAGRATLH"
801..1033
/gene="dp"
3'UTR
ORIGIN
```

```
Query Match 97.8%; Score 1065.4; DB 8; Length 1083;
Best Local Similarity 99.0%; Pred. No. 1.3e-224;
Matches 1072; Conservative 0; Mismatches 11; Indels 0; Gaps 0;
QY 7 GGCACGAGCGCAATGGCGCTCCCGCGCGGAGTGTGTCGCGCGCTACCGCGCACT 66
Db 1 GGCACGAGCGCAATGGCGCTCCCGCGCGGAGTGTGTCGCGCGCTACCGCGCACT 60
QY 67 GGAACCTGACCGCGGTGCACATTTCTGAAAGTCTCCAGTGTCCCGCGCTTCCGGAAGCGG 126
Db 61 GGAACCTGACCGCGGTGCACATTTCTGAAAGTCTCCAGTGTCCCGCGCTTCCGGAAGCGG 120
QY 127 CGGTAATGCGGTCCAAAGGAAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGAA 186
Db 121 CGGTAATGCGGTCCAAAGGAAGGGGCTGTTGACCCGGATAAAGATAGGAAGAGAA 180
QY 187 GCGTGGGCAACCGAGGATCACCGGTTGGGGCTCCCGAGTACAGCAAAATAGTTTGA 246
Db 181 GCGTGGGCAACCGAGGATCACCGGTTGGGGCTCCCGAGTACAGCAAAATAGTTTGA 240
QY 247 GAAAGTTGAAGCCAAAGGAAGAAACAATAGAGTTGCAGACGAAATTTATTTCAGA 306
Db 247 GAAAGTTGAAGCCAAAGGAAGAAACAATAGAGTTGCAGACGAAATTTATTTCAGA 306
```


Itoh, M., Kagawa, I., Kanagawa, S., Katoh, H., Kawai, J., Kishikawa-Hirozane, T., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Oeato, N., Oca, Y., Saichoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Sato, K., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagami-Takeda, Y., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, T., Tomaru, A., Toya, T., Waki, K., Yasunishi, A. and Hayashizaki, Y.

FEATURES
 source
 1. .1174
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /db_xref="taxon:39947"
 /clone="J013105118"

ORIGIN
 Query Match 47.3%; Score 515; DB 8; Length 1174;
 Best Local Similarity 77.4%; Pred. No. 4.1e-103;
 Matches 722; Conservative 0; Mismatches 185; Indels 26; Gaps 7;

118 CGAAGCGCGGTAATCGGTCCAAAGGAAGGGGGCTGTGACCCGGATAAAGATAGGAA 177
 Db
 217 CGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 273
 Qy
 178 GAAGGAGAAGGCTCGCGCAGGAGATCACCGGTGGGGCTCGCGAGTACAGCAAAAT 237
 Db
 274 GAAGGAGAAGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 333
 Qy
 238 AGTTGTGAGAAGTTGAAGCCAAAGGAAGCAACATACATCAGTTCGACGCGAAT 297
 Db
 334 AGTTTCTAAGAAGTTGAGGCCAAAGGAAGCAACATACATCAGTTCGCGATGAT 393
 Qy
 298 TTATTTCAGAGCTGAAGTCCATGGGCACATATTTGGTCAAGGGTTGATCAGAAGATATTAG 357
 Db
 394 TTTTTCGAGAGTGAAGTCCATACGACAGCGTCTGAGTTGATGAGAAGATATTAG 453
 Qy
 358 GCGGAGAGTGTATGATGCTTTCAACGTTCTCATTCGCACTTCGTGTTATTGCAAAAGAAA 417
 Db
 454 GCGGAGGTTATGATGCTTTCAATGTGCTCATTCGCAATTCGTGTTATTGCAAAAGATA 513
 Qy
 418 AAAGGAGATACGTTGGATGGCCCTTTCAATTACAGATATGAATAATAAGAGCTTGA 477
 Db
 514 AAAGAAGATAAAGTGGATGGCCCTTTCTAATTATAGATACGAAAGATACAGAAGTTGGA 573
 Qy
 478 GGAAGTTTCGTAAGAACTCGTCAACAGATAGTAGAAACAAAGAGGCACTCTCTCCAGGAAAT 537
 Db
 574 GGAAGTTTCACAAAGATCTCATCAGAGATCAAGATTAAGNAGAGCTTCTCCAGGAAAT 633
 Qy
 538 CGAAAAACAGTTTGTATGATCTCCTCAAAACATCAAGTTACGTTAACCAACACTGGAAAGCTC 597
 Db
 634 TGAAGAGCAGTTTGTATGATCTCAGATATATACATACGCAACAGGCTAGTCAGAGGCC 693
 Qy
 598 AGCAGAGATGTTAATGGCATCCGCTCCATTCGTTATGTTGTCAGACATCTAGGAAGC 657
 Db
 694 AGCAGAAAGTGTAAATGGCATCCCTTCCTTCTTATGATCAAGACATCCCGAAAGC 753
 Qy
 658 AAGGTTGGAATTCAGATTTTCAGATGATCTCGAAGTTTGGCCATTTTCAGATTCAATGGTGC 717
 Db
 754 AAGGTTGGAATTCAGATTTTCGAGATTCGAAGTTTCGAGGTTTCGACTTCACGGTGC 813
 Qy
 718 ACCATTACATGTCATGATGATCTCTCAATCTTTGAGGGGTAAAGGCTAACACATAGG 777
 Db
 814 ACCATTACCATGTCATGATGATGATGATCAATCTTTGAAGCCATCAGGCGTAA---CAAAG 870
 Qy
 778 AAGAGCTGGCGCGCCACCTTCACT-AGAGACTCAAGATATTACAAATGAATTAAG 836
 Db
 871 AAGAGCTGGCGCTCTCCATTCACCTTTAAGAGGCAAGAATGTGACAAAAACCAATTGAAG 930
 Qy
 837 TG-----TTAGAACTGGCAGCGCGGATTTCTTTTGACAGCTATGTATAGCTTATA--- 886

931 TGTCAAAACTCAAAACATGGCACACAGTTTCTTTTGCACAGTTATGTATAGCTATAGCG 990
 Qy
 887 ----TATCCTCATGAAACATGACCTAGTTTATAGACAGTCTCTCAGGCTTGAAGA- 941
 Db
 991 TACCTTTTCAGTATGGAACATCGACCTAGTTTATAGGACAGTCTCTCAGGCTTTTATAGCAG 1050
 Qy
 942 TTTTAACTCGCAAAATTTGTCCTCTCTTTTGTGCTAGCAGTTATTAGGTCTCAGATAGA 1001
 Db
 1051 ATATGACCTCGCAATTTGCTCTTTTGTGCTAGCAGTTATTAGGTCTCAGATAGA 1106
 Qy
 1002 TGATTTCATATATGCTGCTATGAAACATTTGA 1034
 Db
 1107 TTTTCTTGTGCTGCTGCTATGCTATGATAAACTCA 1139

RESULT 4
 AY224529
 LOCUS
 DEFINITION
 Oryza sativa (japonica cultivar-group) isolate 21044 DP protein
 mRNA, complete cds.
 ACCESSION
 AY224529
 VERSION
 AY224529.1
 GI:29367574
 KEYWORDS
 SOURCE
 ORGANISM
 Oryza sativa (japonica cultivar-group)
 Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzaceae; Oryza.
 REFERENCE
 1 (bases 1 to 885)
 AUTHORS
 Cooper, B., Hutchison, D., Park, S., Guimil, S., Luginbuhl, P.,
 Ellero, C., Goff, S. A. and Glazebrook, J.
 TITLE
 Identification of rice (Oryza sativa) proteins linked to the
 cyclin-mediated regulation of the cell cycle
 JOURNAL
 Plant Mol. Biol. 53 (3), 273-279 (2003)
 MEDLINE
 23111120
 PUBMED
 14750518
 REFERENCE
 2 (bases 1 to 885)
 AUTHORS
 Cooper, B., Hutchison, D., Park, S., Guimil, S., Luginbuhl, P.,
 Ellero, C., Goff, S. A. and Glazebrook, J.
 TITLE
 Direct Submission
 Submitted (27-JAN-2003) Torrey Mesa Research Institute, Syngenta
 Research and Technology, 3115 Merryfield Row, San Diego, CA 92121,
 USA

FEATURES
 source
 Location/Qualifiers
 1. .885
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /isolate="21044"
 /db_xref="taxon:39947"
 /note="derived from yeast two-hybrid experiments"
 CDS
 1. .885
 /codon_start=1
 /product="DP protein"
 /protein_id="AA072649.1"
 /db_xref="GI:29367575"
 /translation="MAPPCGDAAASAAAPLANLLIREGAGLPSRPERVPPRPCTSDS
 FAPISREGDDIPQKTSVLSGGGNAAREEGANNGKEKTKAQRIITGWLRF
 EFSKIYKVKVEAKRTYNEVADEIFAEKLSITONGLEFDEKKNLRRYDAFNVLIAL
 RVIAKDKKEIKWMLNRYEKTOKLEEVHKEITRIKNKKLLOEIKQFDDLONIT
 LRQASORPAESVNGIILPFLIKTSRKARVEIISDSKFAKDFDFNGAPFTMHDDVS
 ILEAIRNKGRAGLSIHP"

ORIGIN
 Query Match 40.6%; Score 442.2; DB 8; Length 885;
 Best Local Similarity 79.8%; Pred. No. 4.6e-87;
 Matches 548; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

118 CGAAGCGCGGTAATCGGTCCAAAGGAAGGGGGCTGTGACCCGGATAAAGATAGGAA 177
 Db
 204 CGCGCGCGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 260
 Qy
 178 GAAGGAGAAGGCTCGCGCAGGAGATCACCGGTGGGGCTCGCGAGTACAGCAAAAT 237


```

KEYWORDS      FLI: CDNA.
SOURCE         Arabidopsis thaliana (thale cress)
ORGANISM       Arabidopsis thaliana
REFERENCE      1 (bases 1 to 879)
AUTHORS        Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
                Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Heuan, V.W.,
                Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
                Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
                Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
                Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
                Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and
                Ecker, J.R.
TITLE          Arabidopsis ORF clones
JOURNAL        Unpublished
REFERENCE      2 (bases 1 to 879)
AUTHORS        Kim, C.J., Chen, H., Cheuk, R., Shinn, P., Bowser, L., Carninci, P.,
                Chan, M.M., Chang, C.H., Dale, J.M., Hayashizaki, Y., Heuan, V.W.,
                Ishida, J., Jones, T., Kamiya, A., Karlin-Neumann, G., Kawai, J.,
                Lam, B., Lee, J.M., Lin, J., Miranda, M., Narusaka, M., Nguyen, M.,
                Onodera, C.S., Palm, C.J., Quach, H.L., Sakurai, T., Satou, M., Seki, M.,
                Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C., Yamada, K.,
                Yu, G., Yuan, S., Shinzaki, K., Davis, R.W., Theologis, A. and
                Ecker, J.R.
TITLE          Direct Submission
JOURNAL        Submitted (14-MAR-2003) Salk Institute Genomic Analysis Laboratory
                (SIGnAL), Plant Biology Laboratory, The Salk Institute for
                Biological Studies, 10010 N. Torrey Pines Road, La Jolla, CA 92037,
                USA
COMMENT        RIKEN Genomic Sciences Center (GSC) members carried out the
                collection and clustering of RAFL cDNAs (RAFL cDNA: "RIKEN
                Arabidopsis Full-Length cDNA") : Seki, M., Narusaka, M., Ishida, J.,
                Satou, M., Kamiya, A., Sakurai, T., Carninci, P., Kawai, J.,
                Hayashizaki, Y. and Shinzaki, K.

The Salk, Stanford, PEGC (SSP) Consortium members constructed and
sequenced the pUNI (ORF) clones using the RAFL cDNAs: Kim, C.J.,
Chen, H., Cheuk, R., Shinn, P., Bowser, L., Chan, M.M., Chang, C.M.,
Dale, J.M., Heuan, V.W., Jones, T., Karlin-Neumann, G., Lam, B.,
Lee, J.M., Lin, J., Miranda, M., Nguyen, M., Onodera, C.S., Palm, C.J.,
Quach, H.L., Southwick, A., Tang, C.C., Toriumi, M., Wong, C., Wu, H.C.,
Yamada, K., Yu, G., Yuan, S., Davis, R.W., Theologis, A., and
Ecker, J.R.

Kim, C.J. (SSP/Salk) and Seki, M. (RIKEN GSC) contributed equally to
this work. Shinzaki, K. (RIKEN GSC) and Ecker, J.R. (SSP/Salk)
contributed equally to this work as PIs.

FEATURES      source
              location/Qualifiers
              1..879
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /db_xref="taxon:3702"
                /chromosome="5"
                /clone="U60005"
                /ecotype="Columbia"
                /note="This clone is in pUNI 51"
              1..879
                /code="putative DP-2 transcription factor"
                /codon_start=1
                /product="At5g02470"
                /protein_id="AA063350.1"
                /db_xref="GI:28950853"
                /translation="MSMEMELFVPEKQHPVSVBEKTPVRRLIVDDSEIGSEKK
                GQSTGGRLQFQVMVYCKLEKITYKEVADEIISDFATIKONAEKPLNNEYNE
                KNERRVYDALNFMALDIARDKKEIRWGLPTCKDVEVKMDKNVSSVOKKA
                AFLKELREKVSLESLMSRNQEMVVKQCPAEGFTLPFLILETNPHAVIEISQDQ
                LVHLDFNSTPFSVHDDAYILIKLMQEQKQNRVSSSSSTTHQSHSSAHSSSSCIAS
                GTSGPVCNWSGSDTR"

CDS
              14.1%; Score 153.6; DB 8; Length 879;

ORIGIN
              Query Match
              14.1%; Score 153.6; DB 8; Length 879;

Best Local Similarity 55.6%; Pred. No. 2.3e-23;
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

QY 131 AATCGCGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGAGAGGCT 190
DB 73 ACTCCAGTGAAGAGAAATTTGATTGTTGATGATGATCTCTGAAATTCGATCAGAGAGAAA 132
QY 191 GCGGACCGAGGATCACCGTTTGGGGCTCCGGAGTACAGCAAAATAGTTTGTGAGAAA 250
DB 133 GGGCAATCAAGAACTTCTGGAGCGGGCTTCGTCATTCAGTTTGTGTTGTTGTCAGAG 192
QY 251 GTTGAAGCCAAAGAGAGAACCAACATACATCAATGAGGTTGCAGCAGAAATTTATTTCAGAGCTG 310
DB 193 TTGGAAGCCCAAGAGATAACTTACTTACAGGAGGTTGCAGCAGAAATTTATTTCAGATTTT 252
QY 311 AAGTCCAT-----GGCACAATATTGGTCAAGGGTTTGTATGAGAAGAA 352
DB 253 GCCACAATTAAGCAAAACGCAGAGAGGCTTTTGAATGAAATGAGTACATGAGAAGAAC 312
QY 353 ATTAGGGGAGAGTGTATGATGCTTTCACAGCTTCTCATTCGACTTCGTGTTATTGCAAAA 412
DB 313 ATAAGCGGAGAGTCTACGATGCGCTCAATGTGTTTCATGCGGTTGGATATATTATTCGAAG 372
QY 413 GAAAAAAGGAGATACGTTGATGGGCTTTTCAAAATTACAGATATGAAAAATAAAGAA 472
DB 373 GATAAAGGAATCCGTTGGAAAGACTTCCTTATACCTGCAAAAGGATGGGAAGAA 432
QY 473 CTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTCCTCCAG 532
DB 433 GTCAAGATGATCGTAATAAGTTATGACAGTGTCAAAAGAGGCTGCTTTCTTTAAA 492
QY 533 GAAATCGAAAAACAGTTTGATGATCTCCAAAACATCAAGTACGTACCAA---ACACTG 589
DB 493 GAGTTGAGAGAAAAGGCTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT 552
QY 590 GAAAGCTCAGCAGAGAATTTAATGGATCGGCTTCCTTATGTTGGTCAAGACATCT 649
DB 553 GTGAAGACTCAAGCCCGCAGAGAGATTACCTTACCTTACCTTACTTACTTGAGACAAAC 612
QY 650 AGGAAAGCAAGGGTGGAAATTTGAGATTTTCAGATGACTCGAAGTTTGCCTTTCGAGTTC 709
DB 613 CCTCAGCAGTAGTCGAAATTCGAGATTTCTGAAGATATGCAACTTGTACACCTCGACTTC 672
QY 710 AATGGTGCCACCATTCACATTCGATGATGAT 739
DB 673 AATAGCACACCTTTCTCGGTCCATGATGAT 702

RESULT 8
ATH319027
LOCUS         Arabidopsis thaliana partial mRNA for E2F dimerisation partner
DEFINITION   protein (dp2a gene).
ACCESSION    AJ319027
VERSION      AJ319027.1 GI:18447785
KEYWORDS     qp2a gene; E2F dimerisation partner protein.
SOURCE       Arabidopsis thaliana (thale cress)
ORGANISM     Arabidopsis thaliana
              Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
              Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
              rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE    1
AUTHORS      Rossignol, P.
TITLE        E2F family transcription factors: AtE2F-a and AtDP-a, induce
              Arabidopsis leaf cells to re-enter S phase
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 960)
AUTHORS      Bergounioux, C.
TITLE        Direct Submission
JOURNAL      Submitted (29-AUG-2001) Bergounioux C., UMR 8618, Cnrs, IBP bat 630
              Université Paris-Sud, 91405 Orsay, FRANCE
FEATURES     source
              location/Qualifiers
              1..960

```

Db 751 AATAGCACACCTTTCTCGGTCCATGATGAT 780

RESULT 9
 AX449297
 LOCUS AX449297 1114 bp DNA linear PAT 03-JUL-2002
 DEFINITION Sequence 16 from Patent WO0185946.
 ACCESSION AX449297
 VERSION AX449297.1 GI:21698044
 KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1.
 REFERENCE Inze D., Boudolf V., de Veylder L., Acosta, J.A. and Magyar, Z.
 AUTHORS Nucleic acid molecules encoding plant cell cycle proteins and uses
 TITLE therefor
 JOURNAL Patent: WO 0185946-A 16 15-NOV-2001;
 CROPDESIGN N.V. (BE)
 FEATURES Location/Qualifiers
 source 1..1114
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN

Query Match 14.1%; Score 153.6; DB 6; Length 1114;
 Best Local Similarity 55.6%; Pred. No. 2.4e-23;
 Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

QY 131 AATGCGGTCCAAAGGAAGGGGGTGTGTGACCCGGATAAGATAGGAAGAGGAGGAGCGCT 190
 DB 86 ACTCCAGTGNAGAGGAATGTATGTTGATGATGATCTGAAATTGGATCAGAGAAGAAA 145

QY 191 GCGGCACCGAGGATCACCGGTTCGGGGCTCCGGAGGTACAGCAAAATAGTTTGTGAGAAA 250
 DB 146 GGGCAATCAAGAACTTCTGGAGGCGGCTTCGTCAATTCAAGTGTATGTTTGTGAGAAG 205

QY 251 GTTGAAGCCAAAGGAAGAACATACATAGAGTTGCAGACCAAAATTTATTTCAGAGCTG 310
 DB 206 TTGGAAGCCAAAGGAATAACTACTTCAAGAGAGGTTGCAGACGAAAATTAATTTTCAGATTTT 265

QY 311 AAGTCCAT-----GGCACATATTGTTCAAGGTTTGTATGAGAAGAAT 352
 DB 266 GCCACATTAAGCAAAACGCAGAGAGCGCTTTGAAATGAATAGTACATGAGAGAAGAC 325

QY 353 ATTAGCGGAGAGTGATGATGCTTTCAACGGTTCTCATTTGCACCTTCGTGTTATTGCAAAA 412
 DB 326 ATAAGCGGAGAGCTACGATGCGCTCAATGTGTTTCATGCGTGGGATATTAATTTGCAAGG 385

QY 413 GAAAAAAGAGATACGGTGGATGGGCGCTTCAAATTAACAGATATGAAAAATAAAGAAG 472
 DB 386 GATAAAAAGGAAATCCGGTGGAAAGGACTTCCTATTACCTGCAAAAAGGATGTGGAGAA 445

QY 473 CTTGAGGAAGTTTCGTAAAGAACTCGTCAACAGATTAGGAACAAGAGGCACCTCCTCCAG 532
 DB 446 GTCAGATGATGCTTAATGAAGTTATGACGAGTGTGCAAAAGAGGCTGCTTTTCTTTAAA 505

QY 533 GAAATCGAAAAACAGTTTGTATGATCTCCAAAAATCAAGTTACGTAACCAA---ACACTG 589
 DB 506 GAGTTGAGAGAAAGGCTCAAGTCTTTGAGAGTCTTATGTGCGAGAAATCAAGAGATGGTT 565

QY 590 GAAAGCTCAGCAGAGAAATGTTAATGGATCCGCGCTTCCATTTCTGATTTGGTCAAGACATCT 649
 DB 566 GTGAAGACTCAAGGCCCCAGCAGAGGAATTTACCTTACCATTCTACTTGTGAGACAAAC 625

QY 650 AGGAAAGCAAGGGTGGAAATTTGAGATTTTCAGATGACTTCGAAGTTTCGCCATTTTCGAGTTC 709
 DB 626 CCTCAGCAGTAGTCGAAATTCGAGATTTCTGGAAGATATGCAACTGTGTACACCTCGACTTC 685

QY 710 AATGGTGCACCAATTCACATTTGCATGATGAT 739

Db 161 ACTCCAGTGAAGAGAAATTGATTGTTGATGATGATCTTGAAATTTGGATCAGAGAAGAA 220
QY 191 GCGCACCGAGGATCACCGGTTGGGGCTCCGGGAGTACAGCAAAATAGTTTGTGAGAA 250
Db 221 GGGCAATCAAGAACTTCTGAGCGGGCTTCGTCATTCAGTGTATGGTTGTCAGAG 280
QY 251 GTTGAAGCCAAAGAGAAACAACATACAAATGAGGTTGCAGACGAAATTTATTTCAGAGCTG 310
Db 281 TTGGAAGCCAAAGAGATAACTTACTTACAGAGGTTGCAGACGAAATTTATTTCAGATTTT 340
QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGTTTGGATGAGAAGAT 352
Db 341 GCCACAATTAAGCAAAACGAGAGAGCCCTTGAATGAAATGAGTACAATGAGAAGAC 400
QY 353 ATTAGCGGAGAGTATGATGCTTCAAGCTTCTCAATTCGACTTCGTTGTTATTGCAAAA 412
Db 401 ATAGCGGAGAGTCTAGATGGCTCAATGTTTCAATGGCTGGATATATTTCAGAG 450
QY 413 GAAAAAAGAGATACGGTGGATGGGCTTTTCAAAATTCAGATATGAAAAATAAAGAG 472
Db 461 GATAAAAAGGAAATCCGGTGGAAAGGACTTCTTATTACCTGCAAAAGAGGATGTGAAGAA 520
QY 473 CTTGAGGAAGTTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCTCCAG 532
Db 521 GTCAAGATGATCGTATAAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTTCTTAAA 580
QY 533 GAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTTACGTAAACAA---ACACTG 589
Db 581 GAGTTGAGAGAAAGGTTCTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT 640
QY 590 GAAAGCTCAGCAGAGAAATGTAATGGCATCGGCTTCCATTCGTTATTGGTCAAGACATCT 649
Db 641 GTCAAGACTCAAGCCAGCAGAGGATTTACCTTACCATTCATCTACTTGAACAAAC 700
QY 650 AGGAAGCAAGGTTGGAATGAGATTCAGATGACTCGAAGTTTGCCTTTCGAGTTC 709
Db 701 CCTCAGCAGTAGTCGAAATCGAGATTTCTGAAGATATGCAACTTGTACACCTCGACTTC 760
QY 710 AATGGTGCACCATTCACATTCATGATGAT 739
Db 761 AATAGCACACCTTCTCGTCCATGATGAT 790

RESULT 12
AX712207
LOCUS AX712207 1274 bp DNA linear PAT 11-APR-2003
DEFINITION Sequence 3 from Patent WO03018818.
ACCESSION AX712207
VERSION AX712207.1 GI:29823429
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE de Veylder L., Inze, D., Mironov, V. and Segers, G.
AUTHORS Method and means for modulating plant cell cycle proteins and their
TITLE use in plant cell growth control
JOURNAL Patent: WO 03018818-A 3 06-MAR-2003;
CROPDESIGN N.V. (BE)
FEATURES Location/Qualifiers
source 1..1274
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 14.1%; Score 153.6; DB 6; Length 1274;
Best Local Similarity 55.8%; Pred. No. 2.4e-23;
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;
QY 131 AATGCGGTCCAAAGGAAGGGGCTGTTGACCGGATAAAGATAGGAAGGAGGAGGCT 190

Db 246 ACTCCAGTGAAGAGAAATTGATTGTTGATGATGATCTTGAAATTTGGATCAGAGAAGAA 305
QY 191 GCGCACCGAGGATCACCGGTTGGGGCTCCGGGAGTACAGCAAAATAGTTTGTGAGAA 250
Db 306 GGGCAATCAAGAACTTCTGAGCGGGCTTCGTCATTCAGTGTATGGTTGTCAGAG 365
QY 251 GTTGAAGCCAAAGAGAAACAACATACAAATGAGGTTGCAGACGAAATTTATTTCAGAGCTG 310
Db 366 TTGGAAGCCAAAGAGATAACTTACTTACAGAGGTTGCAGACGAAATTTATTTCAGATTTT 425
QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGTTTGGATGAGAAGAT 352
Db 426 GCCACAATTAAGCAAAACGAGAGAGCCCTTGAATGAAATGAGTACAATGAGAAGAC 485
QY 353 ATTAGCGGAGAGTATGATGCTTCAAGCTTCTCAATTCGACTTCGTTGTTATTGCAAAA 412
Db 486 ATAAGCGGAGAGTCTACGATGGCTCAATGTTGTTCAATGGCTGGATATATTTCAGAG 545
QY 413 GAAAAAAGAGATACGGTGGATGGGCTTTTCAAAATTCAGATATGAAAAATAAAGAG 472
Db 546 GATAAAAAGGAAATCCGGTGGAAAGGACTTCTTATTACCTGCAAAAGAGGATGTGAAGAA 605
QY 473 CTTGAGGAAGTTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCTCCAG 532
Db 606 GTCAAGATGATCGTATAAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTTCTTAAA 665
QY 533 GAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTTACGTAAACAA---ACACTG 589
Db 666 GAGTTGAGAGAAAGGTTCTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT 725
QY 590 GAAAGCTCAGCAGAGAAATGTAATGGCATCGGCTTCCATTCGTTATTGGTCAAGACATCT 649
Db 726 GTGAAGACTCAAGCCAGCAGAGGATTTACCTTACCATTCATCTACTTGAACAAAC 785
QY 650 AGGAAGCAAGGTTGGAATGAGATTCAGATGACTCGAAGTTTGCCTTTCGAGTTC 709
Db 786 CCTCAGCAGTAGTCGAAATCGAGATTTCTGAAGATATGCAACTTGTACACCTCGACTTC 845
QY 710 AATGGTGCACCATTCACATTCATGATGAT 739
Db 846 AATAGCACACCTTCTCGTCCATGATGAT 875

RESULT 13
AX734251
LOCUS AX734251 1274 bp DNA linear PAT 08-MAY-2003
DEFINITION Sequence 3 from Patent WO03025185.
ACCESSION AX734251
VERSION AX734251.1 GI:30513582
KEYWORDS Arabidopsis thaliana (thale cress)
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
REFERENCE Beeckman, T., de Veylder, L., Inze, D., Mironov, V., Broekaert, W.,
AUTHORS Dillen, W. and Frankard, V.
TITLE A method to modify cell number, architecture and yield of plants by
overexpressing the e2f transcription factor
JOURNAL Patent: WO 03025185-A 3 27-MAR-2003;
CROPDESIGN N.V. (BE)
FEATURES Location/Qualifiers
source 1..1274
/organism="Arabidopsis thaliana"
/mol_type="unassigned DNA"
/db_xref="taxon:3702"

ORIGIN
Query Match 14.1%; Score 153.6; DB 6; Length 1274;
Best Local Similarity 55.6%; Pred. No. 2.4e-23;
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

```
Qy 131 AATCGGTCCAAAGAGAGGGGCTGTTGACCCCGATAAAGATAGGAAGAGGAGGCT 190
Db 246 ACTCCAGTGAAGAGAAATGATGTTGATGATGATCTGAAATTTGGATCAGAGAAGAA 305

Qy 191 GCGGCACCGAGGATCACCGGTTGGGGCTCGCGAGTACAGCAAAATAGTTGTGAGAA 250
Db 306 GGGCAATCAGAACTTCTGGAGCGGGCTTCGTCATTCAGTGTATGTTGTGCAAG 365

Qy 251 GTTGAAGCCAAAGGAAGAACACATACAAATGAGGTTGCAGACGAAATTTATTCAGAGCTG 310
Db 366 TTGGAAGCCAAAGGAAGAACACATACAAATGAGGTTGCAGACGAAATTTATTCAGATTT 425

Qy 311 AAGTCCAT-----GGCACAATTTGGTCAGAGGTTGTATGAGAGAAAT 352
Db 426 GCCACAATTAAGCAAAACGCGAGAAGGCTTTGAATGAAATGAGTACAATGAGAGAAC 485

Qy 353 ATTAGCGGAGAGTATGATGCTTCAACGTTCTCATTTGCACCTTCGTTGTTATTCGCAAA 412
Db 486 ATAAGCGGAGAGTCTAGATCGCTCAATGTGTTTCATGGCTTGGATATTTATTCGAAG 545

Qy 413 GAAAAAAGGAGATACCGGTGATGGGCTTTCAAATTCAGATATGAAAAATAAAGAA 472
Db 546 GATAAAAGGAAATCCGGTGGAAAGGACTTCTATTACCTGCAAAAAGGATGTGGAAGAA 605

Qy 473 CTTGAGAGAGTTCGTAAAGAACTCGTCAACAGATTAAGGAAGAGGACCTCTCTCCAG 532
Db 606 GTCAAGATGGAATCGGTGGAAGAGACTTCTTACCTGCAAAAAGGATGTGGAAGAA 665

Qy 533 GAAATCGAAAAACAGTTGATGATCTCCAAACATCAAGTTTACGTAACCAA---ACACTG 589
Db 666 GTCAAGATGGAATCGGTGGAAGAGACTTCTTACCTGCAAAAAGGATGTGGAAGAA 725

Qy 590 GAAAGCTCAGCAGAGAAATGTAATGCGATCGGCTTCCATTCTGTTGTTGTTGTTGTTGTT 649
Db 726 GTGAGACTCAGGCCAGCAGAGAGATTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 785

Qy 650 AGGAAAGCAGGCTGGAATTTGAGATTTTCAGATGACTTCGAAAGTTTGCCTTTCGAGTTTC 709
Db 786 CCTCAGCAGTACTCGAAATCGAGATTTCTGGAAGATATGCAACTTGTACACCTCGACTTC 845

Qy 710 AATGGTGCACCATTCACATTCGATGATGAT 739
Db 846 AATAGCACACCTTCTCGTCCATGATGAT 875
```

```
RESULT 14
LOCUS ATH294531 1274 bp mRNA linear PLN 15-DEC-2000
DEFINITION Arabidopsis thaliana mRNA for DP-like protein (dpa gene).
ACCESSION AJ294531
VERSION AJ294531.1 GI:11125650
KEYWORDS DP-like protein; dpa gene.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1
Magyar, Z., Atanassova, A., De Veylder, L., Rombauts, S. and Inze, D.
Characterization of two distinct DP-related genes from Arabidopsis
thaliana
PFB Lett. 486 (1), 79-87 (2000)
20562805
PUBMED 11108947
REFERENCE 2 (bases 1 to 1274)
AUTHORS Magyar, Z.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2000) Magyar Z., Department Plantgenetica,
Laboratorium voor Genetica, Universiteit Gent, K.L. Ledeganckstraat
35., B-9000, Gent, BELGIUM
Location/Qualifiers
1. .1274
source
```

```
gene
CDS
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/variety="Columbia"
/db_xref="taxon:3702"
1. .1274
/gene="dpa"
174. .1052
/EC_number="6.3.2.19"
/function="cell cycle regulation"
/codon_start=1
/product="DP-like protein"
/protein_id="CAC15483.1"
/db_xref="GI:11125651"
/db_xref="GOA:Q9FNY3"
/db_xref="Uniprot/TREMBL:Q9FNY3"
/translation="MSMEMELFVTPKQHPHSVSEKTPVRKLVDDSEIGSEKK
GOSRTSGGLRQFVSVVCKLEAKITTYKEVADEIIIDPATIKQNAEKPLNENEYNE
KNIRRVYDALNVFMAIDIIARDKEIRWKLPIITCKKQVDEEVMKRNKVMSSVQKKA
APLKEUREKVSLSLMSRNQEMVVKITQGPAGFTLPIIFILLETNPHAVVEISEDMQ
LVHLDNFSTPFPFVHDDAYILKLMQEQBQBNRVSSSSSTTHQSHSSSSSSCIAS
GTSGPVCMNSGSDTR"
```

ORIGIN

```
Query Match 14.1%; Score 153.6; DB 8; Length 1274;
Best local Similarity 55.8%; Pred. No. 2.4e-23;
Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;
Qy 131 AATCGGTCCAAAGAGAGGGGCTGTTGACCCCGATAAAGATAGGAAGAGGAGGCT 190
Db 246 ACTCCAGTGAAGAGAAATGATGTTGATGATGATCTGAAATTTGGATCAGAGAAGAA 305
Qy 191 GCGGCACCGAGGATCACCGGTTGGGGCTCGCGAGTACAGCAAAATAGTTGTGAGAA 250
Db 306 GGGCAATCAGAACTTCTGGAGCGGGCTTCGTCATTCAGTGTATGTTGTGCAAG 365
Qy 251 GTTGAAGCCAAAGGAAGAACACATACAAATGAGGTTGCAGACGAAATTTATTCAGAGCTG 310
Db 366 TTGGAAGCCAAAGGAAGAACACATACAAATGAGGTTGCAGACGAAATTTATTCAGATTT 425
Qy 311 AAGTCCAT-----GGCACAATTTGGTCAAGGTTTGTATGAGAGAAAT 352
Db 426 GCCACAATTAAGCAAAACGCGAGAAGGCTTTGAATGAAATGAGTACAATGAGAGAAC 485
Qy 353 ATTAGCGGAGAGTATGATGCTTCAACGTTCTCATTCAGCTTCGTTGTTATTCGAA 412
Db 486 ATAAGCGGAGAGTCTACGATCGCTCAATGTGTTTCATGCGCTTGGATATTTATTCGAAG 545
Qy 413 GAAAAAAGGAGATACCGGTGATGGGCTTTCAAATTCAGATATGAAAAATAAAGAA 472
Db 546 GATAAAAGGAAATCGGTGGAAGAGACTTCTTACCTGCAAAAAGGATGTGGAAGAA 605
Qy 473 CTTGAGAGAGTTCGTAAAGAACTCGTCAACAGATTAAGGAAGAGGACCTCTCTCCAG 532
Db 606 GTCAAGATGGAATCGGTGGAAGAGACTTCTTACCTGCAAAAAGGATGTGGAAGAA 665
Qy 533 GAAATCGAAAAACAGTTGATGATCTCCAAACATCAAGTTTACGTAACCAA---ACACTG 589
Db 666 GTTGAAGCCAAAGGAGTCTCAAGTCTTGAAGTCTTATGTCGAGAAATCAAGAGATGGTT 725
Qy 590 GAAAGCTCAGCAGAGAAATGTAATGCGATCGGCTTCCATTCTGTTGTTGTTGTTGTT 649
Db 726 GTGAGACTCAGGCCAGCAGAGAGATTTACCTTACCTTACCTTACCTTACCTTACCTTACCT 785
Qy 650 AGGAAAGCAGGCTGGAATTTGAGATTTTCAGATGACTTCGAAAGTTTGCCTTTCGAGTTTC 709
Db 786 CCTCAGCAGTACTCGAAATCGAGATTTCTGGAAGATATGCAACTTGTACACCTCGACTTC 845
Qy 710 AATGGTGCACCATTCACATTCGATGATGAT 739
Db 846 AATAGCACACCTTCTCGTCCATGATGAT 875
```

RESULT 15
AX100710
LOCUS AX100710 153 bp DNA linear PAT 10-APR-2001
DEFINITION Sequence 7 from Patent WO0121644.
ACCESSION AX100710
VERSION AX100710.1 GI:13619658
KEYWORDS
SOURCE Triticum monococcum
ORGANISM Triticum monococcum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Pooideae; Triticeae; Triticum.
REFERENCE 1
AUTHORS Gutierrez-Armenta, C.C. and Ramirez-Parra, E.C.
TITLE Wheat dp proteins and uses thereof
JOURNAL Patent: WO 0121644-A 7 29-MAR-2001;
CONSEJO SUPERIOR DE INVESTIGACIONES CIENTIFICAS (ES)
FEATURES
source Location/Qualifiers
1..153
/organism="Triticum monococcum"
/mol_type="unassigned DNA"
/db_xref="taxon:4568"
<1..>153
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC36469.1"
/db_xref="GI:13619659"
/translation="DKDRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADE
IYSELKS"
ORIGIN
Query Match 14.0%; Score 153; DB 6; Length 153;
Best Local Similarity 100.0%; Pred. No. 2.6e-23;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 GATAAGATAGGAAGAGGAGGCTGCGGACCGGAGGATCACCGTTGGGGGCTCCGC 223
Db 1 GATAAGATAGGAAGAGGAGGCTGCGGACCGGAGGATCACCGTTGGGGGCTCCGC 60
QY 224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGAGAACACATCAATGAG 283
Db 61 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGAGAACACATCAATGAG 120
QY 284 GTTCGACGAAATTTATTTCAGAGCTGAAGTCC 316
Db 121 GTTCGACGAAATTTATTTCAGAGCTGAAGTCC 153

Search completed: February 6, 2005, 14:35:10
Job time : 5151 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:23:21 ; Search time 697 Seconds
(without alignments)

9249.065 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcgacgagcgca.....aaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004as:*
- 13: Geneseq2004bs:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1084.2	99.6	1089	4	Aaf80144
2	442.2	40.6	885	12	ADQ36842
3	218	20.0	548	10	ADK71795
4	218	20.0	548	12	ADQ63159
5	168	15.4	168	4	Aaf80149
6	168	15.4	168	4	Aaf80148
7	157.6	14.5	1443	12	ADQ62852
8	155.2	14.3	1267	12	ADQ63084
9	155.2	14.3	1267	12	ADQ62144
10	153.6	14.1	1114	6	AAS96287
11	153.6	14.1	1189	6	AAS96319
12	153.6	14.1	1274	10	ADH59559
13	153.6	14.1	1274	10	ACC45100
14	153.4	14.1	1153	12	ADQ63626
15	153	14.0	153	4	Aaf80147
16	151.2	13.9	156	4	Aaf80146
17	150.8	13.8	1158	12	ADQ63082
18	150.8	13.8	1158	12	ADQ62142
19	150.8	13.8	1618	3	AAC39795
20	150	13.8	150	4	Aaf80150

21 138.2 12.7 900 3 AAS59702 DNA encod
22 138.2 12.7 1245 12 ADO63363
23 136.6 12.5 1193 2 AAZ34580 Corn DP-2
24 136.2 12.5 1140 12 ADQ36854
25 132.2 12.1 1041 12 ADQ36844 Cell prol
26 127.8 11.7 962 12 ADO63216
27 123 11.3 1317 12 ADO62851
28 118.6 10.9 1131 6 AAS96304 Arabidops
29 115 10.6 980 12 ADO63362
30 104.6 9.6 425 6 AAS96415
31 102.8 9.4 642 13 ADR65325
32 99.2 9.1 2968 11 ACN89896
33 98.8 9.1 2320 11 ADI32130 Human cDN
34 98.8 9.1 4359 12 ADN05729
35 97 8.9 1266 6 ABL65866 Lung canc
36 97 8.9 1396 5 AAS92244 DNA encod
37 97 8.9 2783 13 ACN38014 Tumour-as
38 97 8.9 2968 5 ABV26838 Human pro
39 97 8.9 2968 5 ABV20992
40 95.2 8.7 665 2 AAZ34577 Impatiens
41 93.6 8.6 870 6 AAS96410 Arabidops
42 93.6 8.6 1442 6 AAS96332 Arabidops
43 88 8.1 1832 2 AAV72862 Caenorhab
44 88 8.1 2203 4 ABL08071 Drosophil
45 85 7.8 358 3 AAC56435 Pinus rad

ALIGNMENTS

RESULT 1
AAF80144
ID AAF80144 standard; DNA, 1089 BP.
AC AAF80144;

XX 11-JUN-2001 (first entry)

DE Nucleotide sequence of a wheat E2F-dimerisation partner (DP) protein.

XX E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
KW ss.

OS Triticum monococcum.

XX Key Location/Qualifiers

FT CDS 20..805

FT /*tag= /a

FT /product= "E2F-dimerisation partner (DP) protein"

XX WO200121644-A2.

XX 29-MAR-2001.

XX 25-SEP-2000; 2000WO-EP009325.

XX 24-SEP-1999; 99ES-00002127.

XX 11-NOV-1999; 99ES-00002474.

XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

XX Gutierrez-Armenta C, Ramirez-Parra E;

XX WPI; 2001-257972/28.

DR P-PSDB; AAB67762.

XX New isolated, enriched, cell free and/or recombinant nucleic acid useful
PT for e.g. altering cell proliferation characteristic such as to alter
PT plant cell, organ or tissue size.

PS Claim 13; Fig 1; 77pp; English.

CC The present sequence encodes a E2F-dimerisation partner (DP) protein. The
CC protein acts as a plant E2F transcription factor. E2F and DP are two
CC proteins that hetero-dimerise to form an active transcription factor that
CC regulates G1 to S phase of the cell cycle, and later, the expression of
CC genes required for S-phase progression. E2F and retinoblastoma protein
CC also interact as a hetero-dimer in cells to suppress certain genes. This
CC repression involves binding of the retinoblastoma protein to the E2F-DP
CC dimer that is in turn bound to sites on DNA through the E2F DNA binding
CC domain. DP proteins can be modulated to alter plant cell, organ or tissue
CC shape, and particularly to alter cell proliferation characteristic such
CC as to alter plant cell, organ or tissue size
XX
SQ Sequence 1089 BP; 367 A; 215 C; 261 G; 246 T; 0 U; 0 Other;

Query Match 99.6%; Score 1084.2; DB 4; Length 1089;
Best Local Similarity 99.7%; Pred. No. 5.2e-207;
Matches 1086; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 GAATTCGGCAGCGCCCAATGGCGCTCCCGCGGCGAGCTGCTCGCGCGCTACCGC 60
DB 1 GAATTCGGCAGCGCCCAATGGCGCTCCCGCGGCGAGCTGCTCGCGCGCTACCGC 60
QY 61 CGCACTGGACTGACCGCGGTGCACATTCCTCGAAGCTTCAGTGTCCCGCGCTCCCGA 120
DB 61 CGCACTGGACTGACCGCGGTGCACATTCCTCGAAGCTTCAGTGTCCCGCGCTCCCGA 120
QY 121 AGCGGCGGTAATCGGTTCCAAAGGAAGGGGGCTTTGACCCCGGATAAAGATAGGAAGAA 180
DB 121 AGCGGCGGTAATCGGTTCCAAAGGAAGGGGGCTTTGACCCCGGATAAAGATAGGAAGAA 180
QY 181 GGAGAAGGCTCGCGCAGCGAGGATCACGGTGTGGGGCTCGCGAGTACAGCAAAATAGT 240
DB 181 GGAGAAGGCTCGCGCAGCGAGGATCACGGTGTGGGGCTCGCGAGTACAGCAAAATAGT 240
QY 241 TTGTGAGAAAGTTCAAGCAGCAAGGAAGCAACATCAATGAGGTTGCGAGTACAGCAAAATTTA 300
DB 241 TTGTGAGAAAGTTGAGCCAAAGGAAGCAACATCAATGAGGTTGCGAGTACAGCAAAATTTA 300
QY 301 TTCAGAGCTGAAGTCCATGGCAGCATATTTGGTCAAGGGTTGATGAGAAGAAATATTAGCG 360
DB 301 TTCAGAGCTGAAGTCCATGGCAGCATATTTGGTCAAGGGTTGATGAGAAGAAATATTAGCG 360
QY 361 GAGAGTGTATGATGCTTTCAACGTTCTCATTGCACTTCGTTTATTCGAAAGAAAAA 420
DB 361 GAGAGTGTATGATGCTTTCAACGTTCTCATTGCACTTCGTTTATTCGAAAGAAAAA 420
QY 421 GGAGATACGGTGGATGGCCCTTTCAAATTCAGATATGAAAAATAAAGAGCTTGAGGA 480
DB 421 GGAGATACGGTGGATGGCCCTTTCAAATTCAGATATGAAAAATAAAGAGCTTGAGGA 480
QY 481 AGTTTCGTAAGAACTCGTCAACAGATAGGAACAAGAGGCACTCTCCAGGAAATCGA 540
DB 481 AGTTTCGTAAGAACTCGTCAACAGATAGGAACAAGAGGCACTCTCCAGGAAATCGA 540
QY 541 AAAACAGTTTTCATGATCTCCTCAAAACATCAAGTACGTAACCAACATCGGAACTCAGC 600
DB 541 AAAACAGTTTTCATGATCTCCTCAAAACATCAAGTACGTAACCAACATCGGAACTCAGC 600
QY 601 AGAGAATGTTAATGGCATCCGCTTTCCATTCGATTTGGTCAAGACATCTAGGAAGCAAG 660
DB 601 AGAGAATGTTAATGGCATCCGCTTTCCATTCGATTTGGTCAAGACATCTAGGAAGCAAG 660
QY 661 GGTGGAATTCAGATTTTCAGATGACTCGAAGTTTGGCCCATTTTCAGTTCAATGTGCACC 720
DB 661 GGTGGAATTCAGATTTTCAGATGACTCGAAGTTTGGCCCATTTTCAGTTCAATGTGCACC 720
QY 721 ATTCACATTCGATGATCTCTCAATCTTGGGGGTAAAGGCTAACACATAGGAAG 780
DB 721 ATTCACATTCGATGATCTCTCAATCTTGGGGGTAAAGGCTAACACATAGGAAG 780
QY 781 AGCTGGCGCGCCACCTTCACTAGAGACTCAAGAAATATTAATGAATTAAGTGT 840
DB 781 AGCTGGCGCGCCACCTTCACTAGAGACTCAAGAAATATTAATGAATTAAGTGT 840

QY 841 AGAACTGGCACAGCCGATCTTTTGCACAGCTATGTATAGCTATATCTCATGAAA 900
DB 841 AGAACTGGCACAGCCGATCTTTTGCACAGCTATGTATAGCTATATCTCATGAAA 900
QY 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGGCTTGAGAAAGATTTTAACTGCAAAATTTG 960
DB 901 CTTGACCTAGTTTATAGGACAGTCTCTCAGGCTTGAGAAAGATTTTAACTGCAAAATTTG 960
QY 961 TCTCCTTTTGTGCTAGCAGGTTATTAGTCTCAGATAGATTCATATATGTGCTGC 1020
DB 961 TCTCCTTTTGTGCTAGCAGGTTATTAGTCTCAGATAGATTCATATATGTGCTGC 1020
QY 1021 TATGAAAAATTCATAGCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 1080
DB 1021 TATGAAAAATTCATAGCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 1080
QY 1081 AAAAAAATAA 1089
DB 1081 AAAAAAATAA 1089

RESULT 2
ADQ36842

ID ADQ36842 standard; DNA; 885 BP.

XX ADQ36842;

XX AC ADQ36842;

XX DT 07-OCT-2004 (first entry)

XX DE Cell proliferation-related nucleic acid sequence #1.
XX KW cell proliferation related polypeptide; cell proliferation; senescence;
XX KW differentiation; stress response; ds.

XX OS Oryza sativa.

XX PN W02004061122-A2.

XX PD 22-JUL-2004.

XX PF 23-DEC-2003; 2003WO-US041200.

XX PR 26-DEC-2002; 2002US-0436565P.

XX PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX PI Cooper B;

XX DR WPI; 2004-534388/51.

XX CC New nucleic acid molecule encoding a cell proliferation-related
XX CC polypeptide, useful for modulating cell proliferation, senescence,
XX CC differentiation, development, and stress response in plants, and for
XX CC producing enhanced food crops.

XX PS Claim 3; SEQ ID NO 1; 408pp; English.

XX CC The present invention relates to an isolated nucleic acid molecule
XX CC encoding a cell proliferation-related polypeptide. The nucleic acid
XX CC molecule and the encoded polypeptide, and methods are useful for
XX CC modulating cell proliferation, senescence, differentiation, development,
XX CC and stress response in plants, and for producing enhanced food crops. The
XX CC present sequence represents a cell proliferation-related nucleic acid
XX CC sequence. The present sequence is published separately from the main body
XX CC of the specification as EPO data.

XX SQ Sequence 885 BP; 249 A; 203 C; 243 G; 190 T; 0 U; 0 Other;

Query Match 40.6%; Score 442.2; DB 12; Length 885;
Best Local Similarity 79.8%; Pred. No. 8.7e-79;
Matches 548; Conservative 0; Mismatches 133; Indels 6; Gaps 2;

AC ADO63159;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2981 orthologous sequence, SEQ ID 1626.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS *Oryza sativa*.
XX
XX WO2004031349-A2.
XX
XX 15-APR-2004.
XX
XX 18-SEP-2003; 2003WO-US030292.
XX
XX 18-SEP-2002; 2002US-0411837P.
XX
XX 17-DEC-2002; 2002US-0434166P.
XX
XX 24-APR-2003; 2003US-0465809P.
XX
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
XX Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX WPI; 2004-330163/30.
XX
XX New recombinant polynucleotide encoding transcription factor
XX polypeptides, useful for producing transgenic plants with advantageous
XX properties compared to a reference plant.
XX
XX Claim 1; SEQ ID NO 1626; 510pp; English.
XX
XX The present invention relates to novel plant transcription factor
XX proteins (I) and nucleotide sequences (II) (ADO631534-ADO63178). The
XX sequences can be used to produce transgenic plants, which overexpress
XX (II), where the transgenic plant has an altered trait as compared to a
XX non-transgenic plant or wild-type plant. The transgenic plant comprises
XX an altered trait selected from increased tolerance to abiotic stress,
XX increased tolerance to osmotic stress, increased tolerance to cold,
XX increased germination in cold, increased tolerance to heat, increased
XX germination in heat, increased tolerance to freezing conditions,
XX increased tolerance to low nitrogen conditions, increased tolerance to
XX low phosphate conditions, increased tolerance to disease, including
XX fungal disease and particularly Erysiphe, Fusarium and Botrytis,
XX increased tolerance to multiple fungal pathogens, increased resistance to
XX glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
XX increased sensitivity to ACC, altered sugar sensing, increased tolerance
XX to sugars, altered carbon/nitrogen sensing, early flowering, late
XX flowering, altered flower structure, loss of flower determinacy, reduced
XX fertility, altered shoot meristem development, altered branching pattern,
XX altered stem morphology, altered vascular tissue structure, reduced
XX apical dominance, altered trichome density, altered trichome development,
XX altered trichome structure, altered root development, altered shade
XX avoidance, altered seed development, altered seed ripening, altered seed
XX germination, slow growth, fast growth, altered cell differentiation,
XX altered cell proliferation, altered cell expansion, altered phase change,
XX altered senescence, abnormal embryo development, altered programmed cell
XX death, lethality when overexpressed, altered necrosis patterns, increased
XX plant size, increased biomass, large seedlings, dwarfed plants, dark
XX green leaves, change in leaf shape, increased leaf size and mass, light
XX green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
XX altered seed coloration, altered seed size, altered seed shape, large
XX seed, increased leaf wax, increased leaf fatty acids, altered seed oil
XX content, altered seed protein content, altered seedprenyl content,
XX altered leaf prenyl lipid content, increased anthocyanin levels, and
XX decreased anthocyanin levels. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences.

XX
SQ Sequence 548 BP; 164 A; 118 C; 112 G; 154 T; 0 U; 0 Other;
Query Match 20.0%; Score 218; DB 12; Length 548;
Best Local Similarity 71.1%; Pred. No. 4.8e-34;
Matches 391; Conservative 0; Mismatches 140; Indels 19; Gaps 7;
QY 554 GATCTCCAAACATCAAGTTACCTTAACCAACACTGGAAGCTCAGCAGAGATGTTAAT 613
DB 1 GACCTTCCAGAAATATACATTACCAACAGGCTAGTCAGAGGCCA-CAGAAAGTGTAAAT 59
QY 614 GGCATCCGCCTTCC-ATTTCGTATTGGTCAAGACATCTAGGAAAGCAGGCTGGAATTTGA 672
DB 60 GGCATCCTCTTCGGGTTCTTATTGATCAAGACATCCGGAAGAGGCTGGAATTTGA 119
QY 673 GATTTCAGATGACTCGAAGTTTCCCATTTTCGAGTTCAATGGTGCACCATTCATTCATGCA 732
DB 120 GATTCGGAAGATTCAAGTTTTCACGGTTTCGACTTCAACGGTGCACCATTCACCATGCA 179
QY 733 TGATGATCTCTCAATCCTTTGAGGGGTTAAGGCGTAAACAGCATAGGAGAGCTGCCGCGC 792
DB 180 TGATGATGTATCAATCTTGAAGCCATCAGGCGGTAAACAAGAGGAGAGCTGCCCTCTC 239
QY 793 CACCTTTCCT-AGAGACTCAAGAAATATTACAATGAAATTTAAAGTG----TTAGAACTG 847
DB 240 CATTACCTCTTAAGAGGCCAAGAATCTGACAAACCATTTGAAGTGTCAAACTCAAACTG 299
QY 848 GCACAGCGGATTCCTTTTGCACAGCTATGATATGATAGCTATA-----TATCCTCATGAAAA 900
DB 300 GCACCCACAGTTTCTTTTGCACAGTTATGTATAGCTATAGCTTACCTTTTCAGTATGAAA 359
QY 901 CTTGACTAGTTTATAGGACAGTCTCTCAGGCTTGAAGA-TTTTAACTGCAAAATTTT 959
DB 360 CTCGACCTAGTTTATAGGACAGTCTCTCAGGCTTTTTCAGCATATGACCTTCGAAATTT 419
QY 960 GTCTCTCTTTTGTGCTTAGCAGGTTATTTAGTCTTCAGATAGATGATTATATGTCGTG 1019
DB 420 GCCTCTTTTGTGCTTAGCAGGTTA---GTCTGGAATAGTTTTCCTTTCGCTGTGCTG 475
QY 1020 CTATGAAACATTTGATAGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1079
DB 476 TTATGATATAAATCAATTCGCGGTTTCAATTCGAAGCTTTGTACAAAGCTTCATTTGATGAAA 535
QY 1080 AAAAAAATAA 1089
DB 536 AAAAAAATAA 545
RESULT 5
AAF80149
ID AAF80149 standard; DNA; 168 BP.
XX
AC AAF80149;
XX
DT 11-JUN-2001 (first entry)
XX
DE DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
XX
KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
KW ss.
XX
OS *Triticum monococcum*.
XX
XX WO200121644-A2.
XX
XX 29-MAR-2001.
XX
XX 25-SEP-2000; 2000WO-EP009325.
XX
XX 24-SEP-1999; 99ES-00002127.
XX
XX 11-NOV-1999; 99ES-00002474.
XX

PA (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 XX Gutierrez-Armenta C, Ramirez-Parra E;
 PI WPI; 2001-257972/26.
 XX P-PSDB; AAB67766.
 DR New isolated, enriched, cell free and/or recombinant nucleic acid useful
 DR for e.g. altering cell proliferation characteristic such as to alter
 DR plant cell, organ or tissue size.
 XX Disclosure; Page 73-74; 77pp; English.
 PS AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
 XX partner (DP) protein. The protein acts as a plant E2F transcription
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an
 CC active transcription factor that regulates G1 to S phase of the cell
 CC cycle, and later, the expression of genes required for S-phase
 CC progression. E2F and retinoblastoma protein also interact as a hetero-
 CC dimer in cells to suppress certain genes. This repression involves
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
 CC be modulated to alter plant cell, organ or tissue shape, and particularly
 CC to alter cell proliferation characteristic such as to alter plant cell,
 CC organ or tissue size
 XX Sequence 168 BP; 62 A; 38 C; 33 G; 35 T; 0 U; 0 Other;
 SQ Query Match 15.4%; Score 168; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 3.7e-24;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 485 CGTAAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAA 544
 Db 1 CGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAA 60
 QY 545 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACACTGGAAGCTCAGCAGAG 604
 Db 61 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACACTGGAAGCTCAGCAGAG 120
 QY 605 AATGTTAATGGCATCGCCCTTCATTCGTTATGTTGTCGAAGACATCTAGG 652
 Db 121 AATGTTAATGGCATCGCCCTTCATTCGTTATGTTGTCGAAGACATCTAGG 168
 RESULT 6
 AAF80148
 ID AAF80148 standard; DNA; 168 BP.
 XX AAF80148;
 XX 11-JUN-2001 (first entry)
 DT DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
 DE E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
 XX S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
 KW ss.
 XX Triticum monococcum.
 OS WO200121644-A2.
 XX 29-MAR-2001.
 PD 25-SEP-2000; 2000WO-EP009325.
 XX 24-SEP-1999; 99ES-00002127.
 PR 11-NOV-1999; 99ES-00002474.
 XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 PA Gutierrez-Armenta C, Ramirez-Parra E;
 PI WPI; 2001-257972/26.
 XX P-PSDB; AAB67766.
 DR New isolated, enriched, cell free and/or recombinant nucleic acid useful
 DR for e.g. altering cell proliferation characteristic such as to alter
 DR plant cell, organ or tissue size.
 XX Disclosure; Page 74-75; 77pp; English.
 PS AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
 XX partner (DP) protein. The protein acts as a plant E2F transcription
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an
 CC active transcription factor that regulates G1 to S phase of the cell
 CC cycle, and later, the expression of genes required for S-phase
 CC progression. E2F and retinoblastoma protein also interact as a hetero-
 CC dimer in cells to suppress certain genes. This repression involves
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
 CC be modulated to alter plant cell, organ or tissue shape, and particularly
 CC to alter cell proliferation characteristic such as to alter plant cell,
 CC organ or tissue size
 XX Sequence 168 BP; 62 A; 38 C; 33 G; 35 T; 0 U; 0 Other;
 SQ Query Match 15.4%; Score 168; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 3.7e-24;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 485 CGTAAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAA 544
 Db 1 CGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAA 60
 QY 545 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACACTGGAAGCTCAGCAGAG 604
 Db 61 CAGTTTGATGATCTCCAAAAACATCAAGTTACGTAACCAACACACTGGAAGCTCAGCAGAG 120
 QY 605 AATGTTAATGGCATCGCCCTTCATTCGTTATGTTGTCGAAGACATCTAGG 652
 Db 121 AATGTTAATGGCATCGCCCTTCATTCGTTATGTTGTCGAAGACATCTAGG 168
 RESULT 6
 AAF80148
 ID AAF80148 standard; DNA; 168 BP.
 XX AAF80148;
 XX 11-JUN-2001 (first entry)
 DT DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
 DE E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
 XX S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
 KW ss.
 XX Triticum monococcum.
 OS WO200121644-A2.
 XX 29-MAR-2001.
 PD 25-SEP-2000; 2000WO-EP009325.
 XX 24-SEP-1999; 99ES-00002127.
 PR 11-NOV-1999; 99ES-00002474.
 XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
 PA Gutierrez-Armenta C, Ramirez-Parra E;
 PI WPI; 2001-257972/26.
 XX P-PSDB; AAB67766.
 DR New isolated, enriched, cell free and/or recombinant nucleic acid useful
 DR for e.g. altering cell proliferation characteristic such as to alter
 DR plant cell, organ or tissue size.
 XX Disclosure; Page 73-74; 77pp; English.
 PS AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
 XX partner (DP) protein. The protein acts as a plant E2F transcription
 CC factor. E2F and DP are two proteins that hetero-dimerise to form an
 CC active transcription factor that regulates G1 to S phase of the cell
 CC cycle, and later, the expression of genes required for S-phase
 CC progression. E2F and retinoblastoma protein also interact as a hetero-
 CC dimer in cells to suppress certain genes. This repression involves
 CC binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
 CC bound to sites on DNA through the E2F DNA binding domain. DP proteins can
 CC be modulated to alter plant cell, organ or tissue shape, and particularly
 CC to alter cell proliferation characteristic such as to alter plant cell,
 CC organ or tissue size
 XX Sequence 168 BP; 59 A; 19 C; 44 G; 46 T; 0 U; 0 Other;
 SQ Query Match 15.4%; Score 168; DB 4; Length 168;
 Best Local Similarity 100.0%; Pred. No. 3.7e-24;
 Matches 168; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 317 ATGCGACATATTGGTCAAGGGTTTGATGAGAAGAATATTAGCGGAGAGTGTATGATGCT 376
 Db 1 ATGCGACATATTGGTCAAGGGTTTGATGAGAAGAATATTAGCGGAGAGTGTATGATGCT 60
 QY 377 TTCAACGGTTCTCATTTGCACTTCCTGTTATTGCAAAAGAAAAAGGAGATACGGTGGATG 436
 Db 61 TTCAACGGTTCTCATTTGCACTTCCTGTTATTGCAAAAGAAAAAGGAGATACGGTGGATG 120
 QY 437 GGCCTTTCAATTTACAGATATGAAATAAAGAAAGCTTCGAGGAAGTT 484
 Db 121 GGCCTTTCAATTTACAGATATGAAATAAAGAAAGCTTCGAGGAAGTT 168
 RESULT 7
 ADO62852
 ID ADO62852 standard; DNA; 1443 BP.
 XX ADO62852;
 XX 15-JUL-2004 (first entry)
 DT Transcription factor G2981/2982 orthologous sequence, SEQ ID 1319.
 DE Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 XX osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.
 XX Glycine max.
 OS WO2004031349-A2.
 XX 15-APR-2004.
 PD 18-SEP-2003; 2003WO-US030292.
 XX 18-SEP-2002; 2002US-0411837P.
 PR 17-DEC-2002; 2002US-0434166P.
 PR 24-APR-2003; 2003US-0465809P.
 XX (MEND-) MENDEL BIOTECHNOLOGY INC.
 PA Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
 PI Riechmann JL, Haake V, Dubell AN, Keddie JS, Sherman BK;

CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,
 CC altered cell proliferation, altered cell expansion, altered phase change,
 CC altered senescence, abnormal embryo development, altered programmed cell
 CC death, lethality when overexpressed, altered necrosis patterns, increased
 CC plant size, increased biomass, large seedlings, dwarfed plants, dark
 CC green leaves, change in leaf shape, increased leaf size and mass, light
 CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
 CC altered seed coloration, altered seed size, altered seed shape, large
 CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
 CC content, altered seed protein content, altered seedprenyl content,
 CC altered leaf prenyl lipid content, increased anthocyanin levels, and
 CC decreased anthocyanin levels. Note: The sequence data for this patent did
 CC not form part of the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1267 BP; 420 A; 225 C; 267 G; 355 T; 0 U; 0 Other;

Query Match 14.3%; Score 155.2; DB 12; Length 1267;
 Best Local Similarity 55.7%; Pred. No. 2e-21;
 Matches 351; Conservative 0; Mismatches 258; Indels 21; Gaps 2;

QY 131 AATCGCGTCCAAAGGAGGGGGCTGTGACCGGATAAAGATAGGAAGGAGAGGCT 190
 DB 239 ACTCCAGTGAAGAGAAATGATGTGTGATGATGATCTGAAATGGTCAGAGAGAAA 298
 QY 191 GCGGACCGAGATACCGGTTGGGGCTCCGCGAGTACAGCAAAATAGTTGTGAGAA 250
 DB 299 GGGCAATCAAGAACTTCTGGAGCGGGCTTCGTCAATTCACTGTTATGTTTCTCAGAAG 358
 QY 251 GTTGAGCCAAAGAGAACCAATACATACATGAGTTGCGACGAAATTTATTCAGAGCTG 310
 DB 359 TTGGAAGCCAAAGAGATACTACTTCAAGGAGGTTGCGACGAAATTTATTCAGATTTT 418
 QY 311 AAGTCCAT-----GGCACATATTGTTCAAGGTTTGTATGAGAGAAAT 352
 DB 419 GCCACAATTAAAGCAAAACGAGAGAGCCTTGGAATGAATAGTACATGAGAGAAC 478
 QY 353 ATTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTTGCACTTCGTTGTTATGCAAAA 412
 DB 479 ATAAGCGGAGAGTCTACGATGCGTCAATGTGTTTCATGCGTGTGATATTATTCAGAG 538
 QY 413 GAAAAAAGAGATACGGTGGATGGGCTTTCAAATACAGATATGAAAAAATAAAGAG 472
 DB 539 GATAAAAGGAATCCGGTGGAAAGACCTCTTATACCTGCAAAAAGGATGTGGAAGAA 598
 QY 473 CTTGAGGAAGTGTGAAGACTCGTCAACAGATTAGGACAGAGCACTCTCCAG 532
 DB 599 GTCAGATGATGTAATAAGTTATGAGCGTGTGCAAAAGAGGCTGCTTTCTTAA 658
 QY 533 GAAATCGAAAAACAGTTGATGATCTCCAAAACATCAAGTTAGCTAACCAA---ACACTG 589
 DB 659 GAGTTGAGAGAAAGTCTCAAGTCTTGAGAGCTTTATGTGGAATCAAGAGATGGTT 718
 QY 590 GAAAGCTCAGCAGAGAAATTTAATGCAATCCGCTTCCATTCGTTATGTGTCAGACATCT 649
 DB 719 GTGAAGACTCAAGGCCCCAGAGAGGATTTACCTTACCATTCTTACTTGTGAGACAAAC 778

QY 650 AGCAAGACGAGGTGGAAATTCAGATTTCAGATGACTCGAAGTTTGCCATTTTCGAGTTC 709
 DB 779 CCTCAGCAGTAGTTCGAATTCGAGATTTCTGAAGATATGCAACTTGTACACCTGACTTC 838
 QY 710 AATGGTGACCAATTCACATTCGATGATGAT 739
 DB 839 AATAGCACACCTTTCTCGGTCCATGATGAT 868

RESULT 9

ADO62144
 ID ADO62144 standard; DNA; 1267 BP.

XX ADO62144;

DT 15-JUL-2004 (first entry)

Transcription factor G2981 coding sequence, SEQ ID 611.

XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
 KW osmotic stress tolerance; cold tolerance; heat tolerance;
 KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
 KW glyphosate resistance; flowering; fertility; seed development; ds.

XX Arabidopsis thaliana.

PN WO2004031349-A2.

XX 15-APR-2004.

PF 18-SEP-2003; 2003WO-US030292.

PR 18-SEP-2002; 2002US-0411837P.

PR 17-DEC-2002; 2002US-0434166P.

PR 24-APR-2003; 2003US-0465809P.

XX (MEND-) MENDEL BIOTECHNOLOGY INC.

PI Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;

PI Riechmann JU, Haake V, Dubell AN, Reddie JS, Sherman BK;

XX WPI; 2004-330163/30.

DR P-PSDB; ADO62145.

PT New recombinant polynucleotide encoding transcription factor
 PT polypeptides, useful for producing transgenic plants with advantageous
 PT properties compared to a reference plant.

XX Claim 1; SEQ ID NO 611; 510pp; English.

XX The present invention relates to novel plant transcription factor
 CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
 CC sequences can be used to produce transgenic plants, which overexpress
 CC (II), where the transgenic plant has an altered trait as compared to a
 CC non-transgenic plant or wild-type plant. The transgenic plant comprises
 CC an altered trait selected from increased tolerance to abiotic stress,
 CC increased tolerance to osmotic stress, increased tolerance to cold,
 CC increased germination in cold, increased tolerance to heat, increased
 CC germination in heat, increased tolerance to freezing conditions,
 CC increased tolerance to low nitrogen conditions, increased tolerance to
 CC low phosphate conditions, increased tolerance to disease, including
 CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
 CC increased tolerance to multiple fungal pathogens, increased resistance to
 CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
 CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
 CC to sugars, altered carbon/nitrogen sensing, early flowering, late
 CC flowering, altered flower structure, loss of flower determinacy, reduced
 CC fertility, altered shoot meristem development, altered branching pattern,
 CC altered stem morphology, altered vascular tissue structure, reduced
 CC apical dominance, altered trichome density, altered trichome development,
 CC altered trichome structure, altered root development, altered shade
 CC avoidance, altered seed development, altered seed ripening, altered seed
 CC germination, slow growth, fast growth, altered cell differentiation,

CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 1267 BP; 420 A; 225 C; 267 G; 355 T; 0 U; 0 Other;

Query Match 14.3%; Score 155.2; DB 12; Length 1267;

Best Local Similarity 55.7%; Pred. No. 2e-21;

Matches 351; Conservative 0; Mismatches 258; Indels 21; Gaps 2;

QY 131 AATCGCGTCCAAAGGAAGGGCGTGTGACCCGATAAAGATAGGAAGGAAGAGGCT 190
DB 239 ACTCCAGTCAGAGGAATGATTTGTTGATGATCTCAATTCGGTCAGAGAAGAAA 298
QY 191 GCGGCACCGAGATCACCGTGGGGCTCCGAGTACAGCAAAATAGTTGTGAGAAA 250
DB 299 GGGCAATCAAGAACTTCGAGCGCGCTTCGCAATTCAGTTATGTTGTCAGAG 358
QY 251 GTTGAAGCCAAAGCAAGAACCAATACATAGCTTGCAGACGAAATTTATTCAGAGCTG 310
DB 359 TTGGAAGCCAGAGATNACTACTTACAGGAGTTGCAGACAAATTTTTCAGATT 418
QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGGTTGATGAGAGAAAT 352
DB 419 GCCACAATTAAGCAAAACGACAGAGAAGCTTTCAATCAAAATCAGTACAAATGAGAAGAAC 478
QY 353 ATTAGCGGAGAGTATGATGCTTTCACAGTTCTCATTCGACTTCGTGTTATTGCAAAA 412
DB 479 ATAAGCGGAGAGTCTACGATGCGCTCAATGTTTCAAGCGTTGATATTATTCAGG 538
QY 413 GAAAAAAGAGATACGGTGGATGGCGCTTTCAAATTTACAGATATGAAAAAATAAAGAAG 472
DB 539 GATAAAAGGAATCCGTTGGAAGGACCTTCTATTACCTGCAAAAGGATGGAAGAA 598
QY 473 CTTGAGGAAGTTCGTAAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACTCCCTCCAG 532
DB 599 GTCAAGATGGATCGTAATAAGTTATGACGAGTGTGCAAAAGAGGCTGCTTTCTTAA 658
QY 533 GAAATCGAAAAACAGTTTGATGATCTCCAAAACATCAAGTTACGTAAACCAA---ACACTG 589
DB 659 GAGTTGAGAGAAAGGTTCTCAAGTCTTGAGAGTCTTATGTCGAGAAATCAAGAGATGGTT 718
QY 590 GAAAGCTCAGCAGAGAATGTAATGGCATCCGCTTCATTCGTATTTGTTCAAGACATCT 649
DB 719 GTGAAGACTCAAGGCCCCAGAGAGGATTTACCTTACCATTTCTTCTTGAGACAAAC 778
QY 650 AGAAGACAGAGGTGGAATGAGATTGAGATTCAGATGACTGCAAGTTGCCATTTTCGAGTTC 709
DB 779 CCTCAGCAGTAGTCGAAATCGAGATTTCGAGAGATATGCAACTTTGTACACCTCGACTTC 838
QY 710 AATGGTGACCATTTCACTTGCATGATGAT 739
DB 839 AATAGCACACTTTCTCGTCCATGATGAT 868

RESULT 10

AAS96287

ID AAS96287 standard; cDNA; 1114 BP.

XX AAS96287;

XX 26-FEB-2002 (first entry)

XX DE Arabidopsis cDNA encoding partial cell cycle protein CCP16.
XX KW Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;
XX KW plant growth regulator; plant development; abiotic stress; biotic stress;
KW nutrient deprivation; pathogen attack; crop yield.
XX OS Arabidopsis thaliana.
XX FN WO200185946-A2.
XX PD 15-NOV-2001.
XX PF 14-MAY-2001; 2001WO-IB001307.
XX PR 12-MAY-2000; 2000US-0204045P.
XX (CROP-) CROPDESIGN NV.
XX PA Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
XX PI
XX DR WPI; 2002-062249/08.
XX DR P-PSDB; AAU72497.
XX PT New cell cycle protein and nucleic acid molecule encoding it useful for
PT regulating cell cycle progression in plants and for identifying
PT modulators which are useful as herbicides or plant growth regulators.
XX PS Claim 38; Fig 16; 316pp; English.
XX CC The invention relates to a novel cell cycle protein (CCP) and the
CC polynucleotides encoding them. CCP is useful for identifying a compound
CC which modulates the activity of the polypeptide and which binds to the
CC polypeptide and an anti-CCP antibody is useful for detecting the presence
CC of CCP in a sample. A CCP modulator is useful for modulating the cell
CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP
CC nucleic acid and polypeptide molecules are useful as modulating agents in
CC regulating cell cycle progression in plants. CCP is useful to treat
CC disorders characterised by insufficient or excessive production of CCP
CC protein or production of CCP protein forms which have decreased or
CC aberrant activity. Compounds that bind to or modulate the activity of CCP
CC polypeptide are useful as herbicides or plant growth regulators. The
CC polynucleotide is useful for modifying cell fate, plant development,
CC plant morphology, biochemistry and/or physiology, the length of the G1,
CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
CC stimulation or enhancement of cell division, DNA replication, seed set,
CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
CC initiation and/or development, module function, dwarfism in plants,
CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
CC and the anti-CCP antibody are useful in agriculture to modulate the
CC protein levels or activity of a protein involved in the cell cycle due to
CC environmental conditions, including abiotic stress such as cold, nutrient
CC deprivation, heat, drought, salt stress, or biotic stress such as
CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate
CC plant architecture, plant quality traits, plant reproduction and seed
CC development, endoreduplication in storage cells, storage tissues and/or
CC storage organs of plants or its parts. CCP is useful as an immunogen to
CC generate antibodies. CCP protein is useful to screen for naturally
CC occurring CCP substrates. The polynucleotide is useful for expressing CCP
CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to
CC modulate CCP activity. The present sequence encodes a CCP protein of the
XX invention
XX SQ Sequence 1114 BP; 378 A; 197 C; 243 G; 296 T; 0 U; 0 Other;

Query Match 14.1%;

Best Local Similarity 55.6%; Pred. No. 4e-21;

Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;

QY 131 AATCGCGTCCAAAGGAAGGGCGTGTGACCCGATAAAGATAGGAAGGAAGAGGCT 190

DB 86 ACTCCAGTCAGAGAAGGAAATGATTTGTTGATGATCTTCAATTCGATCAGAGAAGAAA 145

QY 191 GCAGCACCAGGATCACCGTTGGGGCTCCGCGAGTACAGCAAAATAGTTTGAGAAA 250
 Db 146 GGGCAATCAAGACTTCTGGAGCGGCTTCGTCAATTCAGTGTATGGTTGTCAGAG 205
 QY 251 GTTGAAGCCAAAGAGAAACAATACATCAATGAGGTTCCAGACGAAATTTATTCAGAGCTG 310
 Db 206 TTGGAAGCCAAAGAGATAACTACTTACAGAGGTTTCAGACGAAATTTATTCAGATTTT 265
 QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGGTTTGATGAGAGAAT 352
 Db 266 GCCACATTAAGCAAAAGCAGAGAGCCCTTGAATGAAATGAGTACAATGAGAAGAAC 325
 QY 353 ATTAGCGGAGAGTATGATGCTTTCAACGTTCTCAATTCGACTTCGTGTTATTGCAAAA 412
 Db 326 ATAAGCGGAGAGTCTACGATCGCTCAATGTGTTCAATGGGTTGGATATTTATTCAGAG 385
 QY 413 GAAAAAAGAGATACGTTGGATGGGCTTTCAAAATACAGATATGAAAAATAAGAG 472
 Db 386 GATAAAAGGAAATCCGGTGGAAAGGACTTCTATTACCTGCAAAAAGGATGTGAAGAA 445
 QY 473 CTTGAGGAGCTTCGTAAGAACTCGTCAACAAGATTAGGAACAGAGGCACTCTCCAG 532
 Db 446 GTCAGATGATCGTAAATAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTTCTTAAA 505
 QY 533 GAAATCGAAAAACAGTTTGATGATCTCAAAACATCAAGTTAGCTTAACCAA---ACACTG 589
 Db 506 GAGTTCAGAGAAAGGCTCAAGTCTTGAGAGCTTATGTCGAGAAATCAAGAGATGGTT 565
 QY 590 GAAAGCTCAGCAGAGAAATTAATGAGCATCCGCTTCATTCGATTTGGTCAAGACATCT 649
 Db 566 GTGAAGACTCAAGGCCAGCAGAGAGATTTACCTTACCAATTCATCTTCTGAGACAAAC 625
 QY 650 AGGAAGCAAGGCTGGAATGAGATTTCAGATGACTCGAAGTTTGCCCATTTGCGATTTC 709
 Db 626 CTTACGCGATAGTCGAAATTCGAGATTTCTGAAGATATGCAACTTGTACACCTCGACTTC 685
 QY 710 AATGGTGCACCATTCACATTTGCATGATGAT 739
 Db 686 AATGACACACCTTCTCGGTCCATGATGAT 715

RESULT 11
 AAS96319
 ID AAS96319 standard; cDNA; 1189 BP.
 XX
 AC AAS96319;
 XX
 DT 26-FEB-2002 (first entry)
 XX
 DE Arabidopsis cDNA encoding cell cycle protein CCP16.
 XX
 KW Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;
 KW plant growth regulator; plant development; abiotic stress; biotic stress;
 KW nutrient deprivation; pathogen attack; crop yield.
 XX
 OS Arabidopsis thaliana.
 XX
 PN W0200185946-A2.
 XX
 PD 15-NOV-2001.
 XX
 PF 14-MAY-2001; 2001WO-1B001307.
 XX
 PR 12-MAY-2000; 2000US-0204045P.
 XX
 PA (CROP-) CROPDESIGN NV.
 XX
 PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
 XX
 DR WPI; 2002-062249/08.
 DR P-PSDB; AAU72529.
 XX

PT New cell cycle protein and nucleic acid molecule encoding it useful for
 PT regulating cell cycle progression in plants and for identifying
 PT modulators which are useful as herbicides or plant growth regulators.
 XX
 PS Claim 38; Fig 16; 316pp; English.
 XX
 CC The invention relates to a novel cell cycle protein (CCP) and the
 CC polynucleotides encoding them. CCP is useful for identifying a compound
 CC which modulates the activity of the polypeptide and which binds to the
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP
 CC nucleic acid and polypeptide molecules are useful as modulating agents in
 CC regulating cell cycle progression in plants. CCP is useful to treat
 CC disorders characterised by insufficient or excessive production of CCP
 CC protein or production of CCP protein forms which have decreased or
 CC aberrant activity. Compounds that bind to or modulate the activity of CCP
 CC polypeptide are useful as herbicides or plant growth regulators. The
 CC polynucleotide is useful for modifying cell fate, plant development,
 CC plant morphology, biochemistry and/or physiology, the length of the G1,
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
 CC stimulation or enhancement of cell division, DNA replication, seed set,
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
 CC initiation and/or development, nodule function, dwarfism in plants,
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
 CC and the anti-CCP antibody are useful in agriculture to modulate the
 CC protein levels or activity of a protein involved in the cell cycle due to
 CC environmental conditions, including abiotic stress such as cold, nutrient
 CC deprivation, heat, drought, salt stress, or biotic stress such as
 CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate
 CC plant architecture, plant quality traits, plant reproduction and seed
 CC development, endoreduplication in storage cells, storage tissues and/or
 CC storage organs of plants or its parts. CCP is useful as an immunogen to
 CC generate antibodies. CCP protein is useful to screen for naturally
 CC occurring CCP substrates. The polynucleotide is useful for expressing CCP
 CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to
 CC modulate CCP activity. The present sequence encodes a CCP protein of the
 CC invention
 XX
 SQ Sequence 1189 BP; 400 A; 202 C; 260 G; 327 T; 0 U; 0 Other;

Query Match 14.1%; Score 153.6; DB 6; Length 1189;
 Best Local Similarity 55.6%; Pred. No. 4.1e-21;
 Matches 350; Conservative 0; Mismatches 259; Indels 21; Gaps 2;
 QY 131 AATGCGGTCCAAAGGAAGGGGGCTGTGACCCCGGATAAGATAGGAAGAGGAGAGCT 190
 Db 161 ACTCCAGTGAGAAGGAAATTTGATTGTTGATGATGATCTGAAATTTGGATCAGAGAAGAA 220
 QY 191 GCGGCACCGAGGATCACCGTTGGGGCTCCGCGAGTACAGCAAAATAGTTTGAGAAA 250
 Db 221 GGCAATCAAGAACTTCTGGAGCGGGCTTCGTCATTCAGTTGTTGTTGTCAGAG 280
 QY 251 GTTGAAGCCAAAGAGAAACAATACATCAATGAGTTGCGAGCGAAATTTATTCAGAGCTG 310
 Db 281 TTGGAAGCCAAAGAGATACTACTTACAGAGAGTTGCGAGCGAAATTTATTCAGATTTT 340
 QY 311 AAGTCCAT-----GGCACATATTGGTCAAGGGTTTGATGAGAAGAAT 352
 Db 341 GCCACATTAAGCAAAACGAGAGAAGCCCTTGAATGAAATGAGTACATGAGAAGAAC 400
 QY 353 ATTAGCGGAGAGTGTATGATGCTTTCAAGGTTCTCATTCGACTTCGTGTTATTGCAAAA 412
 Db 401 ATAAGCGGAGAGTCTACGATCGCTCAATGTGTTTCATGCGCTTGGATATTTATTCAGAG 460
 QY 413 GAAAAAAGGAGATACGTTGGATGGGCTTTTCAAAATACAGATATGAAAAATAAGAG 472
 Db 461 GATAAAAGGAAATCCGGTGGAAAGAGCTTCTTATTCCTGCAAAAGAGATGTGGAAGAA 520
 QY 473 CTTGAGGAAAGTTGTAAGAACTCGTCAACAAGATTAGGAACAGAGGCACTCTCCAG 532
 Db 521 GTCAAGATGGATCGTAAATAAGTTATGAGCAGTGTGCAAAAGAGGCTGCTTTCTTAAA 580

CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1153 BP; 337 A; 225 C; 267 G; 324 T; 0 U; 0 Other;

Query Match 14.1%; Score 153.4; DB 12; Length 1153;
Best Local Similarity 55.1%; Pred. No. 4.5e-21;
Matches 354; Conservative 0; Mismatches 271; Indels 18; Gaps 2;
162 CGGATTAAGATAGGAAGAGGAGGCTCGGCGACCGAGGATCACCGGTTGGGGCTCC 221
169 CTGGGAAAAGAAAGAGCTCAGCGTCTACTGAGGGGATAGAGTGTAGAGGACTCA 228
222 GCGAGTACAGCAAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATG 281
229 GACAGTTTAGCATGAAAGTTGTGAGAAAGTTGGAAGCAAGGAAGAACTACGTAT 288
282 AGTTGACAGCAAAATTTATTCAGAGCTGAAGTCCATGTCACATATTGTCAGG---- 336
289 AGTTGACAGTGAAGTGTAGTCTGAGTTTCTGATGCTCAATAGTTGTCAGGCTCAG 348
337 -----GTTTGATCAGAAAGATATTAGCGCGAGGTGTATGATGCTTCAAGCTTC 386
349 ATCAGAAACATATGATGAGAAAGATCATCAGCAGCAGTCTAGATGCTCTGACGTAC 408
387 TCATTGCACTTCGTTTATTGCAAAAGAAAGGAAGATACGTTGATGGGCTTTCAA 446
409 TTATGGCTATGGATATCATTTCTAAGATATAAAGGAATACAGTGAAGGGATTACCAC 468
447 ATTACAGATATGAAAATAAAGAGCTTGAGGAAGTTCTGTAAGAACTCGTCAACAAGA 506
469 GCATGATGCAATGATATTGAGAGCTTAAGACTGAGCGTCTTAAGTGAAGATAGGA 528
507 TTAGGAACAAGAGGCACTCTCCAGGAAATCGAAACAGTTTGTATGATCTCCAAACA 566
529 TTGAAAAGAAAGCAGCTTATTTAGAAAGAACTTGAAGATCAATATGTAGGGCTTCAA 588
567 TCAAGTTAGTTAACCAACACTGGAAGCTCAGCAGAGAAATGTTA---ATGGCATCCGCC 623
589 TCATAAAACGCAATGATCAGTTGTATGGCTCAGGCAATGCTCTAGTGGTGTGGCTT 648
624 TTCCATTCTGTTTGTCAAGACATCTAGGAAGCAAGGGTGGAAATTTGAGATTTTCAGATG 683
649 TACCGTTTATTTAGTGAGACTCTCTCTCATGCTACAGTGAAGTGAATATACAGAAG 708
684 ACTCGAAGTTGCCATTTTCGAGTTTCAATGGTGACCAATTCACATTCATGATGATCTCT 743
709 ATATGAGCTGTGTGATTTTCGACTTTCAACAGCAGCTCCGTTTGTAGCTGATGACGATAT 768
744 CAATCTTTCAGGGGTGAGGCGTAAACAGCATAGGAAGAGCTGG 786
769 ATATCTCTCAAGCAATGAATTTTTCGGGAAGATCAATGACGG 811

RESULT 15
AAF80147
ID AAF80147 standard; DNA; 153 BP.
XX
AC AAF80147;
XX
DT 11-JUN-2001 (first entry)

XX DE
XX KW
XX KW
XX OS
XX PN
XX PD
XX PF
XX PR
XX PR
XX PA
XX PI
XX DR
XX DR
XX PT
XX PT
XX PS
XX CC
XX CC
XX CC
XX CC
XX CC
XX CC
XX CC
XX CC
XX CC
XX CC
XX CC
XX SQ

DNA fragment encoding a wheat E2F-dimerisation partner (DP) protein.
E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
ss.
Triticum monococcum.
WO200121644-A2.
29-MAR-2001.
25-SEP-2000; 2000WO-EP009325.
24-SEP-1999; 99ES-00002127.
11-NOV-1999; 99ES-00002474.
(CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
Gutierrez-Armenta C, Ramirez-Parra B;
WPI; 2001-257972/26.
P-PSDB; AAB67765.
New isolated, enriched, cell free and/or recombinant nucleic acid useful
for e.g. altering cell proliferation characteristic such as to alter
plant cell, organ or tissue size.
Disclosure; Page 72-73; 77pp; English.
AAF80146-50 represent DNA fragments of a gene encoding a E2F-dimerisation
partner (DP) protein. The protein acts as a plant E2F transcription
factor. E2F and DP are two proteins that hetero-dimerise to form an
active transcription factor that regulates G1 to S phase of the cell
cycle, and later, the expression of genes required for S-phase
progression. E2F and retinoblastoma protein also interact as a hetero-
dimer in cells to suppress certain genes. This repression involves
binding of the retinoblastoma protein to the E2F-DP dimer that is in turn
bound to sites on DNA through the E2F DNA binding domain. DP proteins can
be modulated to alter plant cell, organ or tissue shape, and particularly
to alter cell proliferation characteristic such as to alter plant cell,
organ or tissue size

Sequence 153 BP; 55 A; 25 C; 47 G; 26 T; 0 U; 0 Other;
Query Match 14.0%; Score 153; DB 4; Length 153;
Best Local Similarity 100.0%; Pred. No. 3.6e-21;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 164 GATAAAGATAGGAAGAGGAGGCTGCGGCAACCGAGGATCACCGTTGGGGGCTCCGC 223
DB 1 GATAAAGATAGGAAGAGGAGGAGGCTGCGGCAACCGAGGATCACCGTTGGGGGCTCCGC 60
QY 224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACACATCAATGAG 283
DB 61 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACACATCAATGAG 120
QY 284 GTTGCAGACCAAAATTTATTTCAGAGCTCAAGTCC 316
DB 121 GTTGACAGCAAAATTTATTTCAGAGCTCAAGTCC 153
Search completed: February 6, 2005, 13:05:23
Job time : 702 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:47:57 ; Search time 4249 Seconds
(without alignments)
9755.705 Million cell updates/sec

Title: US-10-088-830-1
Perfect score: 1089
Sequence: 1 gaattcgacgagcgca.....aaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:*

1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	614.4	56.4	649	6	CD882927 F1.111L02
2	421.6	38.7	765	6	CA764996 AF53-Rpf
3	394.2	36.2	454	6	CD892664 G118.121L
4	376.4	34.6	607	2	BE456002 HVSMSG001
5	373.8	34.3	612	7	CR289243 CR289243
6	368.2	33.8	1119	9	CL959517 OSIFCC003
7	286.8	26.3	665	6	CA078334 SCLAM100
8	277	25.4	548	2	BE511883
9	264.2	24.3	634	7	CF484481 POL1_25C
10	237.8	21.8	702	7	CR286882 CR286882
11	218	20.0	548	4	B1802607 H084D10 E
12	160.6	14.7	836	9	CL909692 OA_Aba000
13	152.8	14.0	992	7	CK278661 EST724739
14	150.8	13.8	1403	3	EX831191 Arabidops
15	148.2	13.6	1418	3	EX831357 Arabidops
16	143	13.1	1568	3	EX831265 Arabidops
17	142	13.0	680	6	CA290160 SCAGFL801
18	138.2	12.7	1245	5	AV108383 Zea mays
19	135	12.4	501	5	EX680618 BX680618
20	135	12.4	696	9	CL618454 OR_BA001
21	135	12.4	751	9	CL773898 OR_BA008
22	135	12.4	789	9	CL773941 OR_BA008
23	135	12.4	867	9	CL838498 OR_CBa006
24	131.2	12.0	743	7	CN215980 29828 Sus

25	127.8	11.7	498	1	AI939068	AI939068 sc55912.Y
26	124.6	11.4	583	5	BU579005	BU579005 sat64f10.
27	123	11.3	643	6	C96290	C96290 C96290.Narc
28	122.2	11.2	627	1	AU237921	AU237921 AU237921
29	121.8	11.2	550	4	BG040894	BG040894 NXSI_116
30	121	11.1	870	6	CD573622	CD573622 UCRPT01_0
31	119.8	11.0	783	1	AJ804378	AJ804378 AJ804378
32	119	10.9	577	2	BE020396	BE020396 em43901.Y
33	118.4	10.9	728	4	BG873645	BG873645 MEST8-E05
34	118.2	10.9	852	7	CO116984	CO116984 GR_EB019
35	116	10.7	673	7	CK751888	CK751888 eca01-7cs
36	115.6	10.6	760	4	BM406319	BM406319 EST580646
37	115	10.6	980	4	BG320803	BG320803 Zm04_0190
38	113.2	10.4	706	7	CF482425	CF482425 POL1_6 H0
39	112.8	10.4	645	6	CB917390	CB917390 VVD020607
40	112.8	10.4	646	6	CB919100	CB919100 VVD041C12
41	111.4	10.2	790	8	AQ577328	AQ577328 nxb0090F
42	111	10.2	454	5	BU012590	BU012590 QGJ2F12.Y
43	111	10.2	552	7	CK754643	CK754643 pam01-7ms
44	110.6	10.2	640	6	CB919899	CB919899 VVD056H09
45	110	10.1	733	6	CD813391	CD813391 BN15.019H

ALIGNMENTS

RESULT 1
CD882927
LOCUS F1.111L02F010430 F1 Triticum aestivum cDNA clone F111L02, mRNA
DEFINITION sequence.
ACCESSION CD882927
VERSION CD882927.1 GI:32645392
KEYWORDS EST.
SOURCE Triticum aestivum (bread wheat)
ORGANISM Triticum aestivum
REFERENCE 1 (bases 1 to 649)
AUTHORS Genoplatte.
TITLE Genoplatte, a major partnership french program in plant genomics
JOURNAL Unpublished (2003)
COMMENT Contact: Genoplatte
Genoplatte
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french plant genomics programme 'Genoplatte' (http://www.genoplatte.com and http://genoplatte-info.infobiogen.fr).

FEATURES

source
Location/Qualifiers
1..649
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="F111L02"
/tissue_type="leaf one"
/clone_lib="F1"

ORIGIN

Query Match 56.4%; Score 614.4; DB 6; Length 649;
Best Local Similarity 96.8%; Pred. No. 1.9e-123;
Matches 627; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
QY 290 GACGAATTTATTTCAGAGCTGAAGTCCATGGCCACATATTGGTCAAGGTTTGATGAGAAG 349
DB 2 GATGAATTTATTTCAGAGCTGAAGTCCATGGCCACATATTGGTCAAGGTTTGATGAGAAG 61
QY 350 AATATTAGCGGAGAGTGATGATGCTTTCAACGTTTCTCATGTCACCTTCGTTATTGCA 409
DB 62 AATATTAGCGGAGAGTGATGATGCTTTCAACGTTTCTCATGTCACCTTCGTTATTGCA 121


```

RESULT 3
CD892664
LOCUS
DEFINITION
G118.121J23F010725 G118 Triticum aestivum cDNA clone G118121J23,
mRNA sequence.
CD892664
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Triticum aestivum (bread wheat)
ORGANISM
Triticum aestivum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Triticum.
1 (bases 1 to 454)
TITLE
Genoplante.
JOURNAL
Genoplante, a major partnership french program in plant genomics
COMMENT
Unpublished (2003)
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).
FEATURES
Location/Qualifiers
1..454
/organism="Triticum aestivum"
/mol_type="mRNA"
/cultivar="recital"
/db_xref="taxon:4565"
/clone="G118121J23"
/tissue_type="grain (118 degrees per day after
pollination)"
/clone_lib="G118"

```

```

ORIGIN
Query Match 36.2%; Score 394.2; DB 6; Length 454;
Best Local Similarity 96.7%; Pred. No. 2e-75;
Matches 413; Conservative 0; Mismatches 13; Indels 1; Gaps 1;

QY 608 GTTAATGGCATCGCCCTTCATTCGTTATGGTCAAGACATCTAGGAAAGCAAGGGTGGAA 667
Db 11 GGTATGGCATCGCCCTTCATTCGTTATGGTCAAGACATCTAGGAAAGCAAGGGTGGAA 70

QY 668 ATTGAGATTCAGATCGAGTTCGAGTTTGCCCATTTTCGAGTTCAATGGTGCACCATTCACA 727
Db 71 ATTGAGATTCAGACGACTCAAGTGTGCCCATTTCGAGTTCAATGGTGCACCATTCACA 130

QY 728 TTGCATGATGATCTCTCAATCTCTGAGGGGTAAAGCGGTAAACAGCATAGGAAGAGCTGGC 787
Db 131 TTGCATGATGATCTCTCAATCTCTGAGGGGTAAAGCGGTAAACAGCATAGGAAGAGCTGGC 190

QY 788 CGGCCACCCCTTCACCTAGAGACTCAAGATATTTACAAATGAATATAAAG-TGTTAGAACT 846
Db 191 CGGCCACCCCTTCACCTAGAGACTCAAGATATTTACAAATGAATATAAAGTTGTGCAACT 250

QY 847 GGCACAGCGGATCTTTTGCACAGCTATGTATAGCTATATATCTTCATGAAACTTTGAC 906
Db 251 GGCACAGCGGATCTTTTGCACAGCTATGTATAGCTATATATCTTCATGAAACTTTGAC 310

QY 907 CTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGAGATTTTAACTGCAAAATTTGTCTCTCT 966
Db 311 CTAGTTTATAGGACAGCTCTCTCAGGCTTGAGAGAGATTTTAACTGCAAAATTTGTCTCTCT 370

QY 967 TTTTGTGCTAGCAGCTTATTAGTCTCAGATAGATGATTCATATATGCTGCTATGAA 1026
Db 371 TTTTGTGCTAGCAGCTTATTAGTCTCAGATAGATGATTCATATATGCTGCTATGAA 430

QY 1027 AACATTG 1033
Db 431 AACATTG 437

```

```

RESULT 4
BE456002
LOCUS
DEFINITION
HVSMEG0019D17f Hordeum vulgare pre-anthesis spike EST library
cDNA clone HVSMEG0019D17f, mRNA sequence.
BE456002
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
Hordeum vulgare subsp. vulgare
ORGANISM
Hordeum vulgare subsp. vulgare
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Poideae; Triticeae; Hordeum.
1 (bases 1 to 607)
TITLE
Wing,R., Close,R.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,
Yu,X., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
Fenton,R.D., Close,S.J., Oates,R. and Main,D.
Development of a genetically and physically anchored EST resource
for barley genomics: Morex pre-anthesis spike cDNA library
Unpublished (2001)
COMMENT
On Jul 26, 2000 this sequence version replaced gi:13155022.
Contact: Wing RA
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Tel: 864 656 7288
Fax: 864 656 4293
Email: rwing@clemson.edu
Total hg bases = 141
Seq primer: AATTACCTCTCACTAAGG
High quality sequence stop: 585.
FEATURES
Location/Qualifiers
1..607
/organism="Hordeum vulgare subsp. vulgare"
/mol_type="mRNA"
/cultivar="Morex"
/sub_species="vulgare"
/db_xref="taxon:112509"
/clone="HVSMEG0019D17f"
/tissue_type="pre-anthesis spike"
/lab_host="SOLR"
/clone_lib="Hordeum vulgare pre-anthesis spike EST library
HVCDA0008 (white to yellow anther)"
/notes="Vector: lambdaZAP; Site1: EcoRI; Site2: XhoI;
Plants were grown in the greenhouse at the University of
California, Riverside (Fenton, SJ Close, TJ Close). Whole
spike with awns trimmed were collected at white, green and
yellow anther stages (Fenton). Total RNA was prepared from
each pool, equal quantities of all three RNA pools were
combined, poly(A) RNA was purified from the mixture, one
primary unamplified cDNA library was made, and 1 million
pfu were in vivo excised to give pBluescript SK(-) cDNA
phagmids. These steps were performed in the TJ Close lab
(Choi) at the University of California, Riverside.
Phagmids were plated and picked at the Clemson University
Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins
and Wing) Plasmid DNA preparations, DNA sequencing and
sequence analysis were performed at CUGI (Wing, Yu,
Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence
has been trimmed to remove vector sequence and contains a
minimum of 100 bases of phred value 20 or above. For more
details on library preparation and sequence analysis see
http://www.genome.clemson.edu/projects/barley. To order
this clone see http://www.genome.clemson.edu/orders Also
see Close TJ, Wing R, Kleinhofs A, Wise R (2001)
Genetically and physically anchored EST resources for
barley genomics. Barley Genetics Newsletter 31:29-30.
(http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

```

ORIGIN

Query Match

34.6%; Score 376.4; DB 2; Length 607;

```
Best Local Similarity 87.9%; Pred. No. 1.5e-71;
Matches 457; Conservative 0; Mismatches 56; Indels 7; Gaps 4;

QY 9 CACGAGCGCAATGGCGCTCCCGCGGGGAGCTGCTGGCGGGCTACCGCGCACTGG 68
Db 28 CTCGAACCGAATGGCGCTCCCGCGGAGGTGCTGCTGGCGCGTTCGCGCACTGG 87
QY 69 ACTGACCGCGTGCACATCTCTCGAAGCTTCAGAGTTCGCCCGCTCCCGGAA---GCGG 125
Db 88 ACTGACCGCGTGCACATCTCTCGAAGCTTCAGAGTTCGCCCGTTCGCGTTCGCAACGCGGG 147
QY 126 GCGGTAAATGCGTCCAAAGGAGGGGCTGTGACCGGGATGAAGATAGGAAGGAGA 185
Db 148 GCGGCAATGCGTCCAAAGGAGGAGGCGGTTCACCGGATGAAGATAGGAAGGAGA 207
QY 186 AGCTCGGCGCAGGAGTCAACCGGTTGGGGGCTCCCGGAGTACAGCAAAATAGTTTGTG 245
Db 208 AGCTCGGCGCAGGATCAACCGGTTGGGGGCTCCCGGATGAGCAAAATAGTTTGTG 267
QY 246 AGAAAGTTGAAGCAAGGAAGAACACATACATCAATGAGGTTCGAGAGGAAATTTATTCAG 305
Db 268 AGAAAGTTGAAGCAAGGAAGAACACATACATCAATGAGGTTCGAGAGGAAATTTATTCAG 327
QY 306 AGCTGAAGTCCAGGCGACATATTCGTCAGGGTTGATGAGAGAAATATTAGCGCGAGAG 365
Db 328 AGCTGAAGTCCAGGCGACATATTCGTCAGGGTTGATGAGAGAAATATTAGCGCGAGAG 387
QY 366 TGATGATGCTTTCAAGCTTCTCATTCGACCTTCGTTTATTCGAAAGGAAAGGAGA 425
Db 388 GGTATGATGCTTTCAAGCTTCTCATTCGACCTTCATGATATCGCAAAAGAACAAAGGATA 447
QY 426 TAGCGTGGATGGG-CTTTCAAAATTACAGATATGAAAAA--ATAAGAGCTTGAGGAG 482
Db 448 ATCGCGGATGGGCGCTTTCAAAATACAGAAATGAAAGGCTAAACAAAGTTGGAGGAG 507
QY 483 TT-CGTAAAGAACTCGTCAACAGATTAGCAACAGAGG 521
Db 508 TTCCAAAGAACTCGTCAACAGATTATGCAACAGAGG 547

RESULT 5
CR289243 Oryza sativa library (Han B) Oryza sativa cDNA clone
LOCUS p704a10p5, mRNA sequence.
DEFINITION CR289243 Oryza sativa library (Han B) Oryza sativa cDNA clone
VERSION CR289243.1 GI:44675809
KEYWORDS EST.
SOURCE Rice cDNA EST clone
ORGANISM Oryza sativa
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
1 (bases 1 to 612) Ehrhartoideae; Oryzeae; Oryza.
Han, B., Feng, Q., Huang, Y. C., Ying, K., Li, Y., Guan, J. P., Zhu, J. J.,
Zhao, Q., Hu, X., Liu, Y. L., Wu, J., Yu, Z., Chen, L., Fan, D. L.,
Weng, Q. J., Zhang, L., Lu, X. Q., Yu, S. L., Liu, X. H., Lu, F. I.,
Zhang, Y. J., Lu, Y., Li, C., Li, T., Zhang, Y., Hu, H., Jia, P. X.,
Zhang, L., Lan, L. F., Chen, W., Wu, S. A. and Xue, Y. B.
Rice cDNA EST clone
Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
Clone requests: bhan@ncgr.ac.cn
This is rice cDNA est clone
Web site: http://www.ncgr.ac.cn.
Location/Qualifiers
1. 612
/organism="Oryza sativa"
/mol_type="mRNA"

FEATURES
source
```

```
/db_xref="taxon:4530"
/clone="p704a10p5"
/clone_lib="Oryza sativa library (Han B)"

ORIGIN
Query Match 34.3%; Score 373.8; DB 7; Length 612;
Best Local Similarity 79.7%; Pred. No. 5.7e-71;
Matches 441; Conservative 0; Mismatches 112; Indels 0; Gaps 0;

QY 226 GTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGAGAAACAATACATAGAGT 285
Db 1 GTTCAGCAAGATAGTTTCTAAGAAAGTTGAGCCAAAGAGAAACACATATAATGAGT 60
QY 286 TGCAGACGAAATTTATTCAGAGCTCAAGTCCATGCGACATATTCGTCAGGGTTTGATGA 345
Db 61 TCCGATTAAGATTTTTCGGGAGCTGAAGTCCATTAACGAGAACGGTCTGGAGTTGATGA 120
QY 346 GAAGAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCATTTGCACTTCGTGTTAT 405
Db 121 GAAGAATATTAGCGGAGGTTATATGATGCTTTCAATGTCTCATTTGCAATTCGTGTTAT 180
QY 406 TCCAAAGAAAAAGGAGATACGGTGGATGGGCTTTTCAAAATTACAGATATGAAAAAT 465
Db 181 TCCAAAGAAAAAGGAGATAAAGTGGATGGGCTTTACTAATTATAGATACGAAAGAT 240
QY 466 AAAGAGCTTGAAGAGTTTCGTAAAGAACTCGTCAACAGATTAGGAACAAGAGCACT 525
Db 241 ACAGAGTTTGGAGGAGTTTCAAAAGAACTCATACAGGATCAAGATTAAGAGAGT 300
QY 526 CTCCAGGAAATTCGAAAAACAGTTTGTATGATCTCCAAACATCAAGTTACGTAACCAAC 585
Db 301 TCTCCAGGAAATTTGAAAGCAGTTTGTATGACCTTCAGAATATTACATTAACGCAAC 360
QY 586 ACTGAAAGCTCAGCAGAGAAATGTTAAATGGATCGGCTTCATTCGTTTGGTCAAGAC 645
Db 361 TAGTCAGAGCGCAGCAGAAAGTGAATGGCATCTCTCTTCGCTATTATTGATCAAGAC 420
QY 646 ATCTAGGAAGCAAGGTCGAAATTTGAGATTTTCAGATGACTCGAATTTGCGCCATTTCGA 705
Db 421 ATCCGAAAGCAAGGTTGGAATTTGGAATTTGGAAGATTCCAACTTTGCACGGTTCA 480
QY 706 GTTCAATGTCACCATTCACATTCATGATGATCTCTCAATCTCTTGAAGGGGTAGGCG 765
Db 481 CTTCAACTGTGCAACCAATCACCATGATGATATATCAATCCTTGAATCATCATGCG 540
QY 766 TAACAGCATAGGA 778
Db 541 TACCACCAACCGA 553

RESULT 6
CL959517 Oryza sativa indica cultivar-group)
LOCUS OsIFCC003026 Oryza sativa Express Library Oryza sativa (indica
DEFINITION cultivar-group) genomic, genomic survey sequence.
VERSION CL959517
KEYWORDS CL959517.1 GI:52373702
SOURCE GSS.
ORGANISM Oryza sativa (indica cultivar-group)
REFERENCE Oryza sativa (indica cultivar-group)
AUTHORS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
1 (bases 1 to 1119) Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,
Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,
Wong, G. K. S., Deng, X. W. and Wang, J.
An analysis of transcriptional regulation of the rice genome and
its comparison to Arabidopsis
Unpublished (2004)
Contact: Chen Chen
Department of Bioinformatic
Beijing Institute of Genomics
```


Chinese Academy of Sciences, Beijing 101300, China
Tel: 86-10-80481559
Fax: 86-10-80488676
Email: chenchen@genomics.org.cn
Rice genomic sequence.
Class: exon-trapped.

FEATURES

```
1. 1119
  organism="Oryza sativa (indica cultivar-group)"
  /mol_type="genomic DNA"
  /db_xref="taxon:39946"
  /clone_lib="Oryza sativa Express Library"
  /note="Oryza sativa exon trapped genomic sequence"
```

ORIGIN

Query Match	33.8%;	Score 368.2;	DB 9;	Length 1119;
Best Local Similarity	78.8%;	Pred. No. 1e-69;		
Matches 439;	Conservative 0;	Mismatches 118;	Indels 0;	Gaps 0;
QY	248	AAAGTTCAAGCCAAAGGAACAACATACAATCAGGTTCGACGACGAAATTTATTTCAGAG	307	
DB	562	AACGATGAGTCTATACAGCTCAGGAGAGCTTATAGTTGCCGATGAGATTTTTCGCGAG	621	
QY	308	CTGAAGTCCATGCGCACATATTGGTCAAGGGTTTGATGAGAAGATAATTAGCGGAGAGTG	367	
DB	622	CTGAAGTCCATTACGCGAAACGGTCTCGAGTTTGATGAGAAGATAATTAGCGGAGGGA	681	
QY	368	TATGATGCTTTCAACGTTCTCATTTGCACTTCGTGTTATTGCAAAAGAAAAAAGAGATA	427	
DB	682	TATGATGCTTTCAATGTGCTCATTTGCAATTCGTGTTATTGCAAAAGATAAAAAGGAGATA	741	
QY	428	CGGTGGATGGCCCTTTCAAATTACAGATATGAAAAAATAAAGAAAGCTTGAGAAAGTTTCGT	487	
DB	742	AAGTGGATGGCCCTTACTAATTATAGATACGAAAAGATACAGAAGTTGGAGGAAGTTTCAC	801	
QY	488	AAAGAACTCGTCAACAGAGATTAGGAAACAAGAGGCACTCTCTCCAGGAAATCGAAAAACAG	547	
DB	802	AAAGAACTCATCACCAGGATCAAGATAAAGAAGAGCTTCTCCAGGAAATTTGAAAAGCAG	861	
QY	548	TTTGATGATCTCCAAAAACATCAAGTTACGTAAACCACAAACACTGGAAAGCTCAGCAGAGAAT	607	
DB	862	TTTGATGACCTTCAGAAATTTACATTAAGCACCACCGGCTAGTCAGAGCCAGCAGAAAGT	921	
QY	608	GTTAATGGCATCCGCCCTTCCATTCGTATTGGTCAAGACATCTAGGAAAGCAAGGCTGGAA	667	
DB	922	GTTAATGGCATCCCTCTCTCCGTTCTTATTGATCAAGACATCCCAGAAAGCAAGGCTGGAA	981	
QY	668	ATTGAGATTTTCAGATGACTCGAAGTTTGGCCCATTTTCGAGTTCAATGGTGACCAATTCACA	727	
DB	982	ATTGAGATTTTCGGAAGATTCAAAGTTTTCGCGTTTCGACTTCAACGGGTGACCAATTCACC	1041	
QY	728	TTGCATCATGATCTCTCAATCTTTGAGGGGTGAAGCGTAAACGATAGGAAGAGCTGGC	787	
DB	1042	ATGCATGATGATGATCATCTTTGAAGCCATCAGGCGTAAACAAGAGAGAGCTGGC	1101	
QY	788	CGGCCACCCCTTCACTA	804	
DB	1102	CTCTCAATTCACCCCTTA	1118	

RESULT 7	CA078334	665 bp	linear	EST 23-SEP-2003
LOCUS	CA078334			
DEFINITION	SRLAM1006A02.g AM1 Saccharum officinarum cDNA clone SRLAM1006A02			
	5', mRNA sequence.			
ACCESSION	CA078334			
VERSION	CA078334.1	GI:34930606		
KEYWORDS	EST.			
SOURCE	Saccharum officinarum			
ORGANISM	Saccharum officinarum			
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD clade; Panicoideae; Andropogoneae; Saccharum; Saccharum officinarum			

REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

complex.
1 (bases 1 to 665)
Vettore,A.L., da Silva, F.R., Kemper, E.L. and Arruda, P.
The libraries that made SUCEST
Genet. Mol. Biol. 24 (1-4), 1-7 (2001)
Contact: Arruda P
Centro de Biologia Molecular e Engenharia Genetica
Universidade Estadual de Campinas
Caixa Postal 6010, 13083-970, Campinas SP, Brazil
Tel: 55 19 3788 1137
Fax: 55 19 3788 1089
Email: parruda@unicamp.br
Clone distribution: clone distribution information can be found
through the Brazilian Clone Collection Center (BCCC) at
<http://www.bccccenter.fcav.unesp.br>
Plate: 006 row: A column: 02
Seq primer: T7 Promoter Primer.

FEATURES

```

1..665
/organism="Saccharum officinarum"
/mol_type="mRNA"
/db_xref="taxon:4547"
/clone="SCLAM1006A02"
/lab_host="DH10B"
/clone_lib="AM1"
note="Organ: Apical meristem and tissues surrounding of
mature plants; Vector: pSPort1; Site 1: SalI; Site 2:
NotI; An unidirectional cDNA library generated from
[Apical meristem and tissues surrounding of mature
plants]. cDNA was prepared from poly+ mRNA using
Superscript Plasmid System kit (Invitrogen). The
double-strand cDNAs were fractionated in a sepharose
CU-2B 40cm-columns and fragments sizing between 0.8 and
1.5 Kb were directionally cloned into the vector. Details
of each source of RNA and library construction can be
obtained at http://sucst.lad.ic.unimcamp.br/public"

```

ORIGIN

[illegible]

```

QY 521 GCACTCTCCAGAAATCGAAACAGATTGTGATGATCTCAAAACATCAAGTTACGTAAC 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 561 AAACCTCTCCAGAAATCGAACACAGGTTGATGACCTTCAGAATATCAAGTTCCGCAAC 620
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 581 CAAACACTGGA-AAGCTCAGCAGAGAATGTTAATGCGATCCGGCT 624
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 621 CAGTTACTACAGAAGCCAGCAAGAGTGCGAATGGATCAGCCT 665
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 8
BE511863
LOCUS BE511863
DEFINITION 94606410.y1 946 - tassels primordium prepared by Schmidt lab Zea
mays cDNA, mRNA sequence.
ACCESSION BE511863
VERSION BE511863
KEYWORDS BE511863.1 GI:9733131
SOURCE Zea mays
ORGANISM Zea mays
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
1 (bases 1 to 548)
Maize ESTs from various cDNA libraries sequenced at Stanford
University
Unpublished (1999)
Contact: Walbot V
Department of Biological Sciences
Stanford University
855 California Ave, Palo Alto, CA 94304, USA
Tel: 650 723 2227
Fax: 650 725 8221
Email: walbot@stanford.edu
Plate: 946064 row: A column: 10.

FEATURES
     source
     1..548
         /organism="Zea mays"
         /mol_type="mRNA"
         /cultivar="OH43"
         /db_xref="taxon:4577"
         /tissue_type="tassels"
         /dev_stage="just after the transition from vegetative to
         inflorescence development"
         /lab_host="XLOLR"
         /clone_lib="946 - tassels primordium prepared by Schmidt
         lab"
         /note="Organ: tassels; Vector: HybriZAP; Site 1: EcoRI;
         Site 2: XhoI; George Chuck dissected immature tassels
         between 1mm and 3mm. Sharon Stanfield prepared the cDNA
         library in HybriZAP. Sample insert size range was 350 bp
         to 3 Kb with a 1 Kb average."

ORIGIN
Query Match 25.4%; Score 277; DB 2; Length 548;
Best Local Similarity 73.9%; Pred. No. 7.5e-50;
Matches 397; Conservative 0; Mismatches 125; Indels 15; Gaps 3;

QY 53 GCTACCGCCGCACTGGACCTGACCGCGTGCACATTTCTCGAAGCTTCCAGTGTCCCGCCG 112
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 GCTGGCGGCGCAGCCGAGCTCACCAGCTCGCATCAGCAAGAGCTGGATCCCTCCG 77
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 113 CTTCCCGGAA-----CGCGGCGGTATGCGGTCAAAAGGAGGGGGTGTGACCCG 163
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 CTTCCCGGACGGCGAAGCGTTCGGCGCAATGCGAGTGGAGGAGAGGAGGTGATAAG 137
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 164 GATAAGATAGGAGAGGAGAGGCTGCGGCACCGAGGATCACCGTTGGGGGCTCCGC 223
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 AACAGAGCGGAGGAGGTAGGCTGTGACGAGCGGATCCCGGTGGGGGCTCCGC 197
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 224 GAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGAGAAACAAATACATGAG 283
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

```

Db 198 GAGTTCCAGCAAGATAGTTTCTAAAGGTTTGAGACCAAGGACGGACTACTTTACAGTGAG 257
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 284 GTTGCAGACGAAATTTATTCAGAGCTGAAGTCCATGGGCATATTTGGTCAAGGTTTGAT 343
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 GTTGCAGATGAAATTTGTGGAGAGCTGAAGT---TGACTCTTATTGGTCAAGAGTTTGAT 314
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 344 GAGAAGATATTAGGCGGAGAGTGATGATGCTTTTCAAGCTTCTATTTCACCTTCCTGTT 403
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 315 GAGAAGATATCAGGAGGAGAGTGATGATGCTTTTATGTGCTAATTCGACTACGTGTT 374
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 404 ATTGCAAAAGAAAAAGGAGATACGGTGGATGGCGCTTTCAAATTCAGATATGAAAAA 463
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 375 ATCACAAGAGTAAAAAGAGATAAAATGATGGGCTTCTTAATTTCCAATATGAAAAAG 434
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 464 ATAAAGAGCTTGAGGAAGTTCTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCA 523
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 435 ATA--AAGTTGGAGGAACTCGCAAGAACTCATGATCAGGATTAAAGAACAGAAAAA 491
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
QY 524 CTCCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACATCAAGTTACGTAAC 580
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 492 CTTCTCCAGGAAATGAAGACAGCTTGTATGACCTCCAGATATCAAGTTTCGCAAC 548
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
CF484481
LOCUS CF484481
DEFINITION POLI_25_C08.g1_A002 Pollen Sorghum bicolor cDNA clone
POLI_25_C08_A002 5', mRNA sequence.
ACCESSION CF484481
VERSION CF484481.1 GI:34513350
KEYWORDS EST.
SOURCE Sorghum bicolor (sorghum)
ORGANISM Sorghum bicolor
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
clade; Panicoideae; Andropogoneae; Sorghum.
1 (bases 1 to 634)
Cordonnier-Pratt,M.-M., Suzuki,Y., Sugano,S., Klein,R.R., Liang,C.,
Sun,F., Sullivan,R., Eastman,A. and Pratt,L.H.
An EST database from Sorghum: pollen
Unpublished (2003)
Contact: Cordonnier-Pratt MM
Laboratory for Genomics and Bioinformatics
The University of Georgia, Department of Plant Biology
Plant Sciences Building, Rm. 2502, Athens, GA 30602-7271, USA
Tel: 706 542 1860
Fax: 706 583 0210
Email: mmpratt@uga.edu
Library constructed by Dr. Yutaka Suzuki and Dr. Sumio Sugano in
the Human Genome Center, University of Tokyo Institute of Medical
Science; plant material and RNA prepared at Texas A & M University;
sequencing done in the Laboratory for Genomics and Bioinformatics,
University of Georgia. Sequence ends have been trimmed to exclude
vector and regions below Phred quality 16. Three-prime sequences
are presented as their reverse complement and have been trimmed to
exclude polyA.
Seq primer: Sug5 (CTTCTGCTCTAAAGCTGCG).

FEATURES
     source
     1..634
         /organism="Sorghum bicolor"
         /mol_type="mRNA"
         /cultivar="BTx623"
         /db_xref="taxon:4558"
         /clone="POLI_25_C08_A002"
         /lab_host="DH10B-T1 phage-resistant E. coli"
         /clone_lib="Pollen"
         /note="Organ: Pollen; Vector: pME18S-FL3; Site 1: XhoI;
         Site 2: XhoI; The library was prepared from polyA+ RNA
         from pollen at the late vacuolated-vacuolated stage of
         development. Pollen was harvested from greenhouse-grown
         panicles of sorghum line Btx623. Panicles were removed
         from the flag leaf prior to emergence, when no detectable
         amylase is present in pollen of male-fertile lines. This

```

stage represents pollen collected from anthers about 8-14 days prior to anthesis. Double-stranded cDNA was cloned unidirectionally into different DraIII sites of the pME18S-FL3 vector (5-prime DraIII site is CACGTGTG, 3-prime draIII site is CACCATGTG). XhoI excises the cDNA insert."

```

ORIGIN
  Query Match      24.3%; Score 264.2; DB 7; Length 634;
  Best Local Similarity 70.1%; Pred. No. 4.7e-47;
  Matches 385; Conservative 0; Mismatches 158; Indels 6; Gaps 2;

QY 15 CCCAATGGCGCTCCCGCGGCGAGCTCTCGGCGCTACCGCGCACTCGACTGCA 74
Db 92 CCGTAATGGCGCCCTCGGCGGCGAGCGGCGGCGGCGAGCTCACCAGCAGCGCA 151
QY 75 CCGGCGTGACATCTTCGAAGCTTCAGTGTCTCCCGGCTTCCTCCGGAAGCGGCGGTAATG 134
Db 152 TCAGCGAAGGAGCGGCGCTCCCTCCGCTTCCTCGACGGGGAAGCGTCCGCGCAATG 211
QY 135 CGGTCCAAAGGAAGGGGCTGTGACCCGGATAAAGATAGGAAGGAGAGGCTGCGG 194
Db 212 CGGTGAGGAGGAGGAGGCTGTGATACGAACGAAGAGGCGGAAGGAGGCTGGTA 271
QY 195 CACCGAGGATCACCGGTTGGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAAAAAGTTG 254
Db 272 CGCAGCGGATCGCGGTTGGGGGCTTCGCGAGTTCAGCAAGATAGTTTCTAAAAAGTTG 331
QY 255 AGCCAAAGGAAGAAACAACATCAATGAGGTTCCAGACGAAATTTATTCAGAGCTGAAGT 314
Db 332 AGAGCAAAAGGACGCACTACTTCAATGAGGTTCCAGATGAAATTTGTGGGAGCTGAAG- 390
QY 315 CCATGCGACATATGCTCAAGGTTTGTGATGAGAGATATATTAGCGGAGAGTGTATGATG 374
Db 391 --CTGACTCTTATGTGTCAAGAGTTTGTGATGAGAGATATATCAGGAGAGAGTGTATGATG 448
QY 375 CTTTCAACGTTCTCATTTGCACTTCGTGTATTTCAGAAAGAAAAAGGAGATACGCTGGA 434
Db 449 CTTTAAATGTCTAATCGCACTAGCTGTATTACAAAGAACCAAGAGGAGATAAAGTGA 508
QY 435 TGGGCTTTCAATTCAGATATGAAAAATAAAGAGCTTTGAGGAAGTTTCGTAAGAAC 494
Db 509 TGGGCTTGTTCACTTCCATATGAATACATA---AAGTTGGAGGAAACTCGCAAGAAC 565
QY 495 TCGTCAACAGATTAGGAACAAGGCGCACTCTCCAGGAAATCGAAACACATTTGATG 554
Db 566 TCATGATTATGGTTAAGAACAGAAAAAACTTCTCCAGGAAATTTGAACGACAGTTTGTATG 625
QY 555 ATCTCCAAA 563
Db 626 ACCTCCAGA 634

RESULT 10
LOCUS CR286882
DEFINITION CR286882 Oryza sativa library (Han B) Oryza sativa cDNA clone
ACCESSION p730d01p5, mRNA sequence.
VERSION CR286882
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 702)
AUTHORS Han B., Feng Q., Huang Y.C., Ying X., Li Y., Guan J.P., Zhu J.J.,
Zhao Q., Hu X., Liu Y.L., Wu J., Yu Z., Chen L., Fan D.L.,
Weng Q.J., Zhang B., Lu Y.Q., Yu S.L., Liu X.H., Lu T.T.,
Zhang Y.J., Lu Y., Li C., Li T., Zhang Y., Hu H., Jia P.X.,
Zhang L., Lan L.F., Chen W., Wu S.A. and Xue Y.B.
Rice cDNA EST clone

TITLE

```

JOURNAL COMMENT

Unpublished (2003)
Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China
Email: bhan@ncgr.ac.cn
This is rice cDNA est clone
Clone requests: bhan@ncgr.ac.cn
Web site: http://www.ncgr.ac.cn.
Location/Qualifiers
1. .702
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="p730d01p5"
/clone_lib="Oryza sativa library (Han B)"

FEATURES

source

ORIGIN

```

  Query Match      21.8%; Score 237.8; DB 7; Length 702;
  Best Local Similarity 75.7%; Pred. No. 2.7e-41;
  Matches 373; Conservative 0; Mismatches 112; Indels 8; Gaps 6;

QY 192 CGCACCCGAGGATCACCGTTGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAG 251
Db 26 CGGGGCGCAGAGGATCCCGGTTGGGGCTCCGTGAGTTTCAGCAAGATAGTTTCTAAGAAAG 85
QY 252 TTGAAGCCAAAGGAAGAACACATACATGAGGTTTCAGACGCAAAATTTATTTCAGAGCTGA 311
Db 86 TTGAGGCCAAAGGAAGAACACATATATGAGGTTTCGCGATGAGATTTTTCGCGAGCTGA 145
QY 312 AGTCCATGCGACATATTTGTCAGAGGTTTGTATGAGAGAAATATTAGCGGAGAGTGTATG 371
Db 146 AGTCCATGCGACAGAGGTTTCGAGTTTGTATGAGAGAAATATTAGCGGAGGATATATG 205
QY 372 ATGCTTTCAACGTTCTCATTTGCACTTCGTGTATTTCAGAAAGAAAAAGGAGATACGGT 431
Db 206 ATGCTTTC-ATGTCGCTCATTTGCAATTCGTGTATTTCAGAAAGATAATAAAGTATAAAGT 264
QY 432 CGATGGGCTTTCAATTCAGATATGAAAAATAAAGAGCTTTGAGGAGTTTC-GTAAA 490
Db 265 GGATGGGCTTTACTTAATATAGATACGAAAGATACAGAAAGTTTGGAGGAGTTTCACAAA 324
QY 491 GAACTCGTCAACAAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAA-AAAACAGTT 549
Db 325 GAAATCATCACAGGATCCAGATATAGAAAGAGCTTCTCCAGAAATTTGATACGAGAT 384
QY 550 TGATGATCTCCAAAACATCAAGTTACG-TAACCAAAACACTGGAAAGCTCAGCAGAGATG 608
Db 385 TGATGACCTTCAGATATTACATTAAAGCAAAACCAGGCTAGTACGAGGCCAGGACAAAGTG 444
QY 609 TTAATGGCATCC--GCCITTCATTTCGTATTTGTCAGACA--TCTAGGAAGCAAGGGTG 664
Db 445 TTAATGGGATTCATCTACGTTCTTAATGATCAAGAAATTCGATAAAGCAAGGGTG 504
QY 665 GAAATTGAGATTT 677
Db 505 GAAATTGAAATTT 517

```

RESULT 11

```

LOCUS BI802607
DEFINITION BI802607 H084D10 Endosperm library from Oryza sativa (10 days after
ACCESSION BI802607 anthesis) Oryza sativa cDNA clone H084D10, mRNA sequence.
VERSION BI802607.1 GI:16575311
KEYWORDS EST.
SOURCE Oryza sativa
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 548)

```

AUTHORS Dong, H.T., Li, D.B., Zhuang, X.F., Dai, C.G., Sun, L.X., Pei, Y.X.,
Wu, H.F., Jiang, Y.X., Yu, P.C., Gao, Q.K. and Lou, Y.C.
TITLE A Gene Expression Screen in *Oryza sativa*
JOURNAL Unpublished (2001)
COMMENT Contact: Haitao Dong, Debao Li
Bioinformatics and Gene Network Research Group
Zhejiang University
Kaixuan Road 268#, Hangzhou, Zhejiang, P.R.China
Tel: 0086-571-86892051
Fax: 0086-571-86961525
Email: webmaster@estarray.org, URL: http://www.estarray.org
Seq primer: M13 forward primer.
FEATURES
source
1. 548
/organism="Oryza sativa"
/mol_type="mRNA"
/db_xref="taxon:4530"
/clone="H084D10"
/tissue_type="Endosperm"
/dev_stage="10 days after anthesis"
/clone_lib="Endosperm library from *Oryza sativa* (10 days
after anthesis)"
/note="Vector: pSport2"
ORIGIN
Query Match 20.0%; Score 218; DB 4; Length 548;
Best Local Similarity 71.1%; Pred. No. 5.6e-37;
Matches 391; Conservative 0; Mismatches 140; Indels 19; Gaps 7;
QY 554 GATCTCCAAAACATCAAGTTACGTAAACCAACACTGAAAGCTCAGCAGAGAAATGTTAAT 613
DB 1 GACCTTCCAGAAATATACATATACCAACACAGGCTAGTCAGAGGCCA-CAGAAAGTGTAAAT 59
QY 614 GGCAATCGGCTTCC-ATTGCTATTGTCAGAGCATCTAGAGAAAGAGGGTGAATTTGA 672
DB 60 GGCAATCCTCTCCGCTTCTATTGATCAAGACATCCGAAAGCAAGGGTGAATTTGA 119
QY 673 GATTTCCAGATGACTCCGAAGTTGCCATTTCCAGTTCAATGGTGCCACCATTCACATTGCA 732
DB 120 GATTTCCGAAGATTCGAAGTTTTCACGGTTTCGACTTCAACGGTGCACCATTCACCATGCA 179
QY 733 TGATGATCTCTCAATCCTTGAGGGGTAAAGCGCTAACAGCATAGGAAGAGCTGGCGCGC 792
DB 180 TGATGATGATCAATCCTTGAAGCCATCAGGCGTAAACAAAGGAAGAGCTGGCGCTC 239
QY 793 CACCTTCTACT-AGAGCTCAAGAAATATTACAAATGAATTAAAGTG-----TTAGAACTG 847
DB 240 CATTCACCTTAAAGAGGCAAGAATGTGACAAACCATTTGAAGTGTCAAACTCAAAACTG 299
QY 848 GCACAGCGGATCTTTTGACAGCTATGATAGCTATA-----TATCCTCATGAAA 900
DB 300 GCACACACAGTTCTTTTGACAGCTATGATAGCTATAGCGTACCTTTTCAGTATGAAA 359
QY 901 CTTGACTAGTTTATAGGACAGTCTCTCAGGCTTGAGAAGA-TTTTAACTCGAAATTTT 959
DB 360 CTGCACTAGTTTATAGGACAGTCTCTCAGGCTTTTATAGCATATGACCTTGGAAATTT 419
QY 960 GTCTCCTTTTGTGCTTGTAGGATTAATAGGCTTCAGATAGATGATTCATATATGCTG 1019
DB 420 GCCTCTTTTGTGCTTGTAGGATTAATAGGCTTCAGATAGATGATTCATATATGCTG 475
QY 1020 CTATGAACATTTGATAGCAAAAAAATAAAAAAATAAAAAAATAAAAAAATAAAAAA 1079
DB 476 TTATGATATAAATCAATTCGCGTTTCAATTCGAAGCTTGTGATCAAGCTTCAATTCAT 535
QY 1080 AAAAAAATA 1089
DB 536 AAAAAAATA 545
RESULT 12
CL909692/c 836 bp DNA linear GSS 14-SEP-2004
LOCUS
CL909692
DEFINITION CK278661 992 bp mRNA linear EST 03-AUG-2004
EST724739 potato abiotic stress cDNA library Solanum tuberosum cDNA

DEFINITION OA_Aba0009114.f OA_Aba *Oryza australiensis* genomic clone
ACCESSION OA_Aba0009114 5', genomic survey sequence.
VERSION CL909692
KEYWORDS CL909692.1 GI:52018571
SOURCE GSS.
ORGANISM *Oryza australiensis*
Oryza australiensis
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; *Oryza*.
REFERENCE 1 (bases 1 to 836)
AUTHORS Kim, H., Yu, Y., Stum, D., Yost, D., Rao, K., Luo, M., Jetty, R.,
Kudrna, D., Muller, C., Hatfield, J., Soderlund, C. and Wing, R.
TITLE OMAP Project
JOURNAL Unpublished (2004)
COMMENT Contact: Rod A. Wing
Arizona Genomics Institute
University of Arizona
Forbes Building Room 303, Tucson, AZ 85721-0036, USA
Tel: 520 626 9595
Fax: 520 621 1259
Email: http://genome.arizona.edu
PCR Primers
FORWARD: TAA TAC GAC TCA CTA TAG GG
BACKWARD: CAC TCA TTA GGC ACC CCA
Plate: 0009 row: I column: 14
Seq primer: TAA TAC GAC TCA CTA TAG GG
Class: BAC ends.
FEATURES
Location/Qualifiers
1. 836
/organism="Oryza australiensis"
/mol_type="genomic DNA"
/db_xref="taxon:4532"
/clone="OA_Aba0009114"
/tissue_type="young leaves"
/lab_host="DH10B T1 phage resistant"
/clone_lib="OA_Aba"
/note="Vector: pGIBAC1; Site_1: HindIII; Site_2: HindIII"
ORIGIN
Query Match 14.7%; Score 160.6; DB 9; Length 836;
Best Local Similarity 74.8%; Pred. No. 1.9e-24;
Matches 246; Conservative 0; Mismatches 69; Indels 14; Gaps 3;
QY 714 GTGCACCATTCACATTCATGATGATCTCTCAATCCTCAGGGGTAAAGCGTAACAGCA 773
DB 344 GTGCACCATTCACATTCATGATGATGATCTCTCAATCCTCAGGGGTAAAGCGTAACAGCA 285
QY 774 TAGGAAGAGCTGGCGCGCACCCCTTCACT--AGAGACTCAAGAAATATTACAAATGAAT 831
DB 284 TAGGAAGAGCTGGCGCTCTCCATTCCTTTAAAGAGACACAGAAATGTGACAAAAACAGT 225
QY 832 AAAAGTGTTAGAACTGGCACAGCCGATTTCTTTTGACAGCTATGATAGCTAT-- 885
DB 224 TGAAGTGTCAAAACTGCCACTGCCAGTTCCTTTTGGACAGCTATGATAGCTATAGCGTA 165
QY 886 --ATATCCTCATGAAAACCTTGACCTAGTTTATAGGACAGCTCTCAGGCTTGAGAGATT 943
DB 164 CCATTCAGTAATGAAAACCTCGACCTAGTTTATAGGACAGCTATCTCAGGCTTTTAAACAA 105
QY 944 TTAACCTGCAAAATTTTGTCTCTTTTGTGCTTGTAGCAGGTTATTAGGCTCTCAGATAGT 1003
DB 104 TATGACTCGGAATTTTGGCTCTTTTGTGCTTGTAGCAGGTTA--GTCTGGATAGGTG 49
QY 1004 ATTATATATATGCTGCTGATGAAAACATT 1032
DB 48 TTCTTTCATGCTGCTGCTATGATAAACT 20
RESULT 13
CK278661
LOCUS
DEFINITION CK278661 992 bp mRNA linear EST 03-AUG-2004
EST724739 potato abiotic stress cDNA library Solanum tuberosum cDNA

```

clone POAE738 5' end, mRNA sequence.
CK278661
CK278661.1 GI:39835639
EST.
Solanum tuberosum (potato)
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Solanum.
1 (bases 1 to 992)
Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
Generation of ESTs from abiotic stressed potato tissue
Unpublished (2003)
Other ESTs: EST724740
Contact: Robin Buell
The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from the University of Arizona Genomics
Institute via http://genome.arizona.edu/orders/
Seq primer: ATT TAG GTG ACA CTA TAG.
FEATURES
    source
        1..992
            /organism="Solanum tuberosum"
            /mol_type="mRNA"
            /cultivar="Kennebec"
            /db_xref="taxon:4113"
            /clone="POAE738"
            /tissue_types="abiotic stress treated leaf and root tissue"
            /lab_host="DHI08-Tona"
            /notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
            supplier: Solanum tuberosum var. Kennebec plants were
            grown from cuttings on a 16hr light/8 hr dark cycle at 25
            C for 3-4 weeks. Abiotic stress conditions were applied to
            four separate sets of plants. Set 1 involved saturation of
            the soil with 150 mM NaCl and tissues were harvested at
            following application of the salt stress (leaves: 2hr,
            6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d).
            Set 2 were grown under the standard conditions and then
            were water stressed by withdrawal of further watering
            applications. Drought stressed plants were harvested after
            cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d
            and 5d). Set 3 were grown under the standard conditions
            and then were cold stressed by placement at 4 C. Cold
            stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
            and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
            and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d,
            then were heat stressed by placement at 35 C. Heat
            stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d,
            2d and 4d and heat-stressed roots were harvested at 6 hr,
            12 hr, 1 d, and 4d. RNA was isolated from all tissues and
            equal RNA from each tissue and stress was pooled to
            construct the cDNA library. RNA sample."
ORIGIN
    Query Match 14.0%; Score 152.8; DB 73; Length 992;
    Best Local Similarity 55.5%; Pred. No. 9.9e-23;
    Matches 342; Conservative 0; Mismatches 252; Indels 18; Gaps 2;
QY 162 CGGATAAAGATAGCAAGAGGAGGAGGCTCGCGCACCAGGAGTACCGGGTGGGGCTCC 221
DB 48 CTGGGAAAGAAAGAGCTACCGTGTACTCTGGAGGGGTAGAGTGTAGAGGACTCA 107
QY 222 GCGAGTACGACAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAGCAACATACAATG 281
DB 108 GACAGTTCCAGCATGAAAGTTGTGAGAAAGTTGAAGCCAAAGGAGCAACATACAATG 167
QY 282 AGGTTCCAGCAAGAAATTTATTCAGAGCTGAGTCCATGGCACATATTGTCAGG----- 336
DB 168 AGGTTCTGATGAAGTGTGAGCTGAGTCTTTCTGATGCTACCAATAGTTGTCAGGCCAG 227
QY 337 -----GTTTGTAGAGCAATATTATGAGCGGAGAGTGTATGATGCTTTTCAACGTTTC 386

```

clone POAE738 5' end, mRNA sequence.

CK278661

CK278661.1 GI:39835639

EST.

Solanum tuberosum (potato)

Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamiales; Solanales; Solanaceae; Solanum.

1 (bases 1 to 992)

Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.

Generation of ESTs from abiotic stressed potato tissue

Unpublished (2003)

Other ESTs: EST724740

Contact: Robin Buell

The Institute for Genomic Research

9712 Medical Center Dr, Rockville, MD 20850, USA

Email: potato-array@tigr.org

Clones can be requested from the University of Arizona Genomics Institute via http://genome.arizona.edu/orders/

Seq primer: ATT TAG GTG ACA CTA TAG.

FEATURES

source

1..992

/organism="Solanum tuberosum"

/mol_type="mRNA"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="POAE738"

/tissue_types="abiotic stress treated leaf and root tissue"

/lab_host="DHI08-Tona"

/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI; supplier: Solanum tuberosum var. Kennebec plants were grown from cuttings on a 16hr light/8 hr dark cycle at 25 C for 3-4 weeks. Abiotic stress conditions were applied to four separate sets of plants. Set 1 involved saturation of the soil with 150 mM NaCl and tissues were harvested at following application of the salt stress (leaves: 2hr, 6hr, 12hr, 1d, 2d, and 4d; roots: 2hr, 6hr, 12hr, and 2d). Set 2 were grown under the standard conditions and then were water stressed by withdrawal of further watering applications. Drought stressed plants were harvested after cessation of watering (leaves: 3d, 5d, and 7d; roots: 3d and 5d). Set 3 were grown under the standard conditions and then were cold stressed by placement at 4 C. Cold stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, and 4d and roots were harvested at 2 hr, 6 hr, 12 hr, 1 d, then were heat stressed by placement at 35 C. Heat stressed leaves were harvested at 2 hr, 6 hr, 12 hr, 1 d, 2d and 4d and heat-stressed roots were harvested at 6 hr, 12 hr, 1 d, and 4d. RNA was isolated from all tissues and equal RNA from each tissue and stress was pooled to construct the cDNA library. RNA sample."

ORIGIN

Query Match 14.0%; Score 152.8; DB 73; Length 992;

Best Local Similarity 55.5%; Pred. No. 9.9e-23;

Matches 342; Conservative 0; Mismatches 252; Indels 18; Gaps 2;

QY 162 CGGATAAAGATAGCAAGAGGAGGAGGCTCGCGCACCAGGAGTACCGGGTGGGGCTCC 221

DB 48 CTGGGAAAGAAAGAGCTACCGTGTACTCTGGAGGGGTAGAGTGTAGAGGACTCA 107

QY 222 GCGAGTACGACAAATAGTTGTGAGAAAGTTGAAGCCAAAGGAGCAACATACAATG 281

DB 108 GACAGTTCCAGCATGAAAGTTGTGAGAAAGTTGAAGCCAAAGGAGCAACATACAATG 167

QY 282 AGGTTCCAGCAAGAAATTTATTCAGAGCTGAGTCCATGGCACATATTGTCAGG----- 336

DB 168 AGGTTCTGATGAAGTGTGAGCTGAGTCTTTCTGATGCTACCAATAGTTGTCAGGCCAG 227

QY 337 -----GTTTGTAGAGCAATATTATGAGCGGAGAGTGTATGATGCTTTTCAACGTTTC 386

228 ATCAGAAACAATATGATGAGAGCAATCATCAGACGACGAGTCTAGGATGCTCTGACGTAC 287

387 TCATTGCACTTCGTGTTATTTGCAAAAGAAAAAGAGATACGGTGGATGGCCCTTTCAA 446

288 TTATGGCTATGGATATCATTTCTTAAAGATATAAAGGAAAAATACAGTGGAGGGATTACCAC 347

447 ATTACAGATATGAAAAAATAAGAGAGCTTCGAGGAAGTTCCGTAAGAACTCGTCAACAGA 506

348 GCACTGATGCAATGATATTTGAGAGCTTAAAGACTGAGCGCTTAACTTGAGAAATAGGA 407

507 TTAGGAACCAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTGATGATCTCCAAAACA 566

408 TTCAAAAGAAAGCAGCGCTATTAGAGAACTTGAAGATCAATATGTAGGGCTTCAAAACC 467

567 TCAAGTTACGTAACCAACACTCGAAAGCTCAGCAGAGAAATGTTA---ATGGCATCCGCC 623

468 TCATAAAACGCAATGATGATGATGTTATGCTCAGGCAATGCTCTAGTGGTGGTGGCTT 527

624 TTCATTTCGTATTGTTCAAGACATCTAGGAAGCAAGGGTGGAAATTTGAGATTTTCAGATG 683

528 TACCGTTTATTTAGTGCAGACTCGCTCTCATCAACAGTTGAAGTGAAGTGAATATCAGAAG 587

684 ACTCGAAGTTTGCCCATTTGAGTTCAATGGTGCACCATTCACATTCGATGATGATCTCT 743

588 ATATGAGCTGTGTCATTTTCGACTTCAACAGCACTCGGTTTGAGTACATGACGATAATT 647

744 CAATCTTGAGG 755

648 ATATCTCAAAG 659

RESULT 14

CNS09YMA

LOCUS

DEFINITION

Arabidopsis thaliana Full-length cDNA Complete sequence from clone GSLTUS672A12 of Adult vegetative tissue of strain col-0 of Arabidopsis thaliana (thale cress).

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

JOURNAL

COMMENT

FEATURES		Location/Qualifiers	
source	1. .1403	/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/strain="Col-0"	
		/db_xref="taxon:3702"	
		/clone="GSLTSL672A12"	
gene		/tissue_type="Adult vegetative tissue"	
		/plasmid="PCWVSPORT_6"	
	1. .1403	/gene="At5g03415"	
ORIGIN		Query Match	
		13.8%; Score 150.8; DB 3; Length 1403;	
		Best Local Similarity 55.3%; Pred. No. 2.8e-22;	
		Matches 346; Conservative 0; Mismatches 262; Indels 18; Gaps 2;	
QY	147	AGGGGCTGTTGACCCCGATAGAGTAGGAAGAGGAGGCTGCGGACCGAGGATCA	206
Db	426	AAGAGCTTCTGTTGTTAAGAAAGAGAGGGGACAGCGTGGCTGGTCCAGATAGA	485
QY	207	CCGGTTGGGGGCTCCCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAA	266
Db	486	CTGGAAGAGGACTACGCTCAATTAGTATGAGAGTTGTGAAAAGGTGGAAGCAAGGAA	545
QY	267	GAACAACATACATGAGGTGCGAGCAAAATTTATTCAGAGCT-----GA	311
Db	546	GGACAACCTACATGAGGTGCGAGCGAGCTTGTCTGAATTTGCATTTCCAAATGACG	605
QY	312	AGTCCATGGCACATATTGGTCAAGGTTTGATGAGAAAGATATTAGCGGAGAGTGTATG	371
Db	606	ATGGAACATCCCTGATCAGCAACATGATGATGAGAAACATAGACGAGAGTATATG	665
QY	372	ATGCTTTCAACGTTCTCATTCGACTTCTGTTTATTCGAAAGAAAAAGGAGATCGGT	431
Db	666	ATGCTTTAAACGTCCTCATGGCTATGATATAATACCAAGGATATAAAAGAAATTCAT	725
QY	432	GGATGGCCCTTCAATATACAGATATGAAAAATAAGAGCTTGAGGAGTTCGTAAG	491
Db	726	GGAGAGGTCTTCTCGGCAAGCTTAAGCGACATTAAGAAATTAAGAGACGAACTCT	785
QY	492	AACCTGTCACAAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTG	551
Db	786	CATTAGGACAGAAATTGAGAAGAAACTGCATATTTCCCAAGAACTCGAAGACATATG	845
QY	552	ATGATCTCCAAACATCAAGTTACGTAAACCAACACTGGAAAGCTCAGCAGAGATGTTA	611
Db	846	TAGGCTTTCAGAACTCTGATACAGAGAAATGAGCACTTATATAGCTCAGGAAATGCTCCA	905
QY	612	ATGCG---ATCCGCTTCCATTCGTTATTTGTCAGACATCTAGGAAGCAAGGTTGAA	668
Db	906	GTGCGGTGTGTCTTCTTTTATCTCTGTCAGACTCGTCTCAGCAACAGTAGAAG	965
QY	669	TTGAGATTTCAGATGACTCGAAAGTTTCCCAATTTCCAGTTTCAATGTGCAACCATTCACAT	728
Db	966	TCGAGATATCAGAGATATCGAGCTCGTGCAATTTGATTTCAACAGCACTCCATTTGAGC	1025
QY	729	TGCATGATGATCTCTCAATCCTTGAAG	754
Db	1026	TCACAGCAGCAATTTTGTCTCTCAAG	1051
RESULT 15		CNS09YMC	
		1418 bp mRNA linear HTC 06-FEB-2004	
		Arabidopsis thaliana Full-length cDNA Complete sequence from clone	
		GSLTSL812A03 of Adult vegetative tissue of strain col-0 of	
		Arabidopsis thaliana (thale cress).	
		BX831357	
		BX831357.1 GI:42454999	
		HTC; GSLT cDNA.	
		Arabidopsis thaliana (thale cress)	
		Arabidopsis thaliana	
ORGANISM		Bukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
		Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
		rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
REFERENCE		Castelli, V., Aury, J.M., Jallou, O., Wincker, P., Clepet, C.,	
AUTHORS		Menard, M., Cruaud, C., Quetier, P., Scarpelli, C., Schachter, V.,	
TITLE		Temple, G., Caboche, M., Weissenbach, J., and Salanoubat, M.	
JOURNAL		Whole Genome Sequence Comparisons and 'Full-length' cDNA Sequences:	
AUTHORS		A Combined Approach to Evaluate and Improve Arabidopsis Genome	
TITLE		Annotation	
JOURNAL		Unpublished	
AUTHORS		2 (bases 1 to 1418)	
TITLE		Genoscope.	
JOURNAL		Direct Submission	
AUTHORS		Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :	
TITLE		BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr	
JOURNAL		- Web : www.genoscope.cns.fr)	
COMMENT		The sequences are based on single pass reads.	
		Life Technologies (a division of Invitrogen) members carried out	
		full-length libraries construction : Temple G.	
		Genoscope members carried out sequencing and annotation : Castelli	
		V., Aury J.M., Jallou O., Wincker P., Menard M., Cruaud C.,	
		Schachter V., Weissenbach J., Salanoubat M.	
		URGV INRA : Clepet C., Caboche M.	
		Annotation is based on the June 2003 version of the Arabidopsis	
		genome released by MIPS (Munich Information center for Protein	
		Sequences). 5 prime and 3 prime are assembled with Phrap.	
		http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_Full	
		length	
		http://www.genoscope.cns.fr/cgi-bin/ggb?source=Arabidopsis.	
FEATURES		Location/Qualifiers	
source	1. .1418	/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/strain="Col-0"	
		/db_xref="taxon:3702"	
		/clone="GSLTSL812A03"	
gene		/tissue_type="Adult vegetative tissue"	
		/plasmid="PCWVSPORT_6"	
	1. .1418	/gene="At5g03415"	
ORIGIN		Query Match	
		13.6%; Score 148.2; DB 3; Length 1418;	
		Best Local Similarity 55.0%; Pred. No. 1e-21;	
		Matches 345; Conservative 0; Mismatches 263; Indels 19; Gaps 2;	
QY	147	AGGGGCTTTCACCCCGATAGGATAGGAAGAGGAGGCTGCGGACCGAGGATCA	206
Db	410	AAGGAGCTTCTGTTGTTAAGAAAGAGAGGGGACAGCGTGGCTGGTCCAGATAAGA	469
QY	207	CCGGTTGGGGGCTCCCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAA	266
Db	470	CTGGAAGAGGACTACGCTCAATTTAGTATGAGAGTTTGTGAAAAGGTGGAAGCAAGGAA	529
QY	267	GAACAACATACATGAGTTGCGAGCAAAATTTATTCAGAGCT-----GA	311
Db	530	GGACAACCTTACATGAGTTGCGAGCAAGCTTGTCTGAATTTGCATTTCCAAATAACG	589
QY	312	AGTCCATGGCACATATTGGTCAAGGTTTGTATGAGAAAGATATTAGCGGAGAGTGTATG	371
Db	590	ATGGAACATCCCTGATCAGCAACATGATGAGAAAAACATAAGCAAGAGTATATG	649
QY	372	ATGCTTTCAACGTTCTCATTTGCATTCGTTTATTCGAAAGAAAAAGGAGATACGGT	431
Db	650	ATGCTTTAAACGTCCTCATGGCTATGATATATATCCAAAGGATATAAAAGAAATTCAT	709
QY	432	GGATGGCCCTTTCAAATTTACAGATATGAAAAATAAGAGCTTGAGGAGTTTCGTAAG	491
Db	710	GGAGAGTCTTCTCGGCAAGCTTAAGCGACATTCGAGAAATTAAGACGACGACTCT	769
QY	492	AACCTGTCACAAAGATTAGGAACAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTG	551

```

Db 770 CACTTAGGACAGATTGAGAGAGAAACTGCATATTCCTCAAGAACTGGAAGACAATATG 829
QY 552 ATGATCTCCAAAACATCAAGTTACGTAAACCAACACACTGGAAAGCTCAGCAGAGAAATGTTA 611
Db 830 TAGGCCTTCAGATCTGATACAGAGAAATGAGCACTTATATAGCTCAGGAAATGCTCCCA 889
QY 612 ATGGC----ATCGCCTTCCATTCTGTTGGTCAAGACATCTAGGAAAGCAAGGGTGGAA 667
Db 890 GTGGCGGTGTTGCTCTTCTCTTTTATCTTGTCCAGACTCGTCTCAGCAACAGTAGAA 949
QY 668 ATTGAGATTCAGATGACTCGAAGTTTGCCCAATTCGAGTTCAATGGTGCACCAATTCACA 727
Db 950 GTGGAGATATCAGAGATATGCAGCTCGTGCAATTTGATTCAACAGCACTCCCAATTTGAG 1009
QY 728 TTGCATGATGATCTCTCAATCCTTTGAG 754
Db 1010 CTCCAGCAGACAAATTTGTCTCTCAAG 1036

```

Search completed: February 6, 2005, 15:46:10
Job time : 4259 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 12:25:31 ; Search time 228 Seconds
(without alignments)
7815.374 Million cell updates/sec

Title: US-10-088-830-1

Perfect score: 1089

Sequence: 1 gaattcgacgagcgca.....aaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued Patents NA:*

- 1: /cgn2_6/prodata/1/ina/5A COMB.seq.*
- 2: /cgn2_6/prodata/1/ina/5B COMB.seq.*
- 3: /cgn2_6/prodata/1/ina/6A COMB.seq.*
- 4: /cgn2_6/prodata/1/ina/6B COMB.seq.*
- 5: /cgn2_6/prodata/1/ina/PCTUS COMB.seq.*
- 6: /cgn2_6/prodata/1/ina/backfile1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98.8	9.1	2320	4	US-09-023-655-1456
2	97	8.9	1427	4	US-09-949-016-3349
3	86.2	7.9	1157	2	US-08-723-415B-5
4	86.2	7.9	1157	3	US-09-189-627A-5
5	86.2	7.9	1157	3	US-09-710-861-5
6	85	7.8	358	4	US-09-640-211A-566
7	83.2	7.6	1700	2	US-08-428-131-1
8	83.2	7.6	1700	3	US-08-602-846-1
9	83.2	7.6	1700	3	US-09-078-596-1
10	82.8	7.6	364	4	US-09-640-211A-465
11	81.2	7.5	1154	2	US-08-723-415B-3
12	81.2	7.5	1154	3	US-09-189-627A-3
13	81.2	7.5	1154	3	US-09-710-861-3
14	81.2	7.5	1202	2	US-08-723-415B-7
15	81.2	7.5	1202	3	US-09-189-627A-7
16	81.2	7.5	1202	3	US-09-710-861-7
17	81.2	7.5	1385	2	US-08-723-415B-1
18	81.2	7.5	1385	3	US-09-189-627A-1
19	81.2	7.5	1385	3	US-09-710-861-1
20	74.8	6.9	1440	4	US-09-023-655-1399
21	74.8	6.9	1672	4	US-09-949-016-2937
22	71.6	6.6	601	4	US-09-949-016-120497
C 22	71.6	6.6	601	4	Sequence 120497
C 23	71.6	6.6	193169	4	Sequence 120498
24	69.4	6.4	1461	3	US-08-722-126A-4
25	69.4	6.4	1461	5	PCT-US95-04258-4
26	69.4	6.4	7218	1	US-08-232-463-14
C 27	69.4	6.4			Sequence 14, Appl

28	65.8	6.0	1008	4	US-09-780-641-1	Sequence 1, Appl
29	65.8	6.0	2202	3	US-09-465-558-59	Sequence 95, Appl
30	65.2	6.0	578	3	US-09-602-877A-95	Sequence 95, Appl
31	64.4	5.9	189	4	US-09-621-976-14761	Sequence 14761, A
32	63.8	5.9	1507	3	US-09-453-323-1	Sequence 1, Appl
33	63.8	5.9	1813	3	US-09-071-224-3	Sequence 3, Appl
34	63.6	5.8	1034	4	US-09-311-021-105	Sequence 105, App
35	63.6	5.8	1141	4	US-09-800-729-78	Sequence 78, Appl
36	63.2	5.8	1736	3	US-09-182-816-22	Sequence 22, Appl
C 37	63.2	5.8	1736	3	US-09-182-816-24	Sequence 22, Appl
38	63.2	5.8	1736	3	US-09-471-528-22	Sequence 22, Appl
C 39	63.2	5.8	1736	3	US-09-471-528-24	Sequence 24, Appl
40	63.2	5.8	1736	3	US-09-634-530-22	Sequence 22, Appl
C 41	63.2	5.8	1736	3	US-09-634-530-24	Sequence 24, Appl
42	62.8	5.8	1602	1	US-08-530-950-3	Sequence 3, Appl
43	62.8	5.8	1602	3	US-08-888-429A-3	Sequence 3, Appl
44	62.8	5.8	1602	3	US-09-149-879-3	Sequence 3, Appl
45	62.8	5.8	1602	4	US-09-057-009-3	Sequence 3, Appl

ALIGNMENTS

RESULT 1
US-09-023-655-1456
; Sequence 1456, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; FILING DATE: HEREWITH
; APPLICATION NUMBER: US/09/023,655
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1456:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9604478
; US-09-023-655-1456

Query Match 9.1%; Score 98.8; DB 4; Length 2320;

Best Local Similarity 58.1%; Pred. No. 6.8e-15;
Matches 200; Conservative 0; Mismatches 132; Indels 12; Gaps 1;

QY 236 ATAGTTTGTGAGAAAGTTGAGCCAAAGGAAGAACATACATACATGAGTTGACAGGAA 295
DB 369 AAAGTGTGTGAGAAAGTTCAACGAAAGGTACAAACATCGTACAAATGAAGTCGCTGATGAG 428
QY 296 ATTTATTCAGAGCTG-----AAGTCATGCGACATATTTGGTCAAGGGTTTGTAT 343
DB 429 CTGGTGTACAGAGTTCACCAATTCAAATAACCATTTGGCTGCTGATTTCGCAGGGCTTAATGAT 488
QY 344 GAGAAATATTTAGCGGAGAGTGTATGATGCTTTTCAAGCTTCTCATTCGACACTTCGTGTT 403
DB 489 CAGAAACATTTAGGCGAAGAGTTTATGATGCTTTAAATGTGCTATGCAATGAACATA 548
QY 404 ATTGCAAAAGAAAAAGGAGATACCGTGTGATGGCTTTCAAAATACAGATATGAAAAA 463
DB 549 ATTCAAGGAAAAAGAAATCAAGTGTGATGGCTGCTACCAATTTCTGCTCAGGAA 608
QY 464 ATAAAGAGCTTGAGGAAGTTCTGAAGAACTGTCACAAAGATTAGGAACAGAGGCA 523
DB 609 TGTCAAGATCTGGAGATAGAGAGCGAGCGGATAGAACGGATAAAGCAGAAAGCGGGCC 668
QY 524 CTCTCCAGGAATCGAAAAACAGATTGTGATGATCTCCAAAAACAT 567
DB 669 CAGCTGCAAGAACTTCTCTACAGCAAAATCGCTTTCAAAAACCT 712

RESULT 2
US-09-949-016-3349
; Sequence 3349, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3349
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3349

Query Match 8.9%; Score 97; DB 4; Length 1427;
Best Local Similarity 57.8%; Pred. No. 1.6e-14;
Matches 197; Conservative 0; Mismatches 135; Indels 9; Gaps 1;

QY 236 ATAGTTTGTGAGAAAGTTGAGCCAAAGGAAGAACATACATACATGAGTTGACAGGAA 295
DB 459 AAAGTGTGTGAGAAAGTTCAACGAAAGGTACAAACATCGTACAAATGAAGTCGCTGATGAG 518
QY 296 ATTTATTCAGAGCTG-----AAGTCATGCGACATATTTGGTCAAGGGTTTGTATGAG 346
DB 519 CTGGTGTACAGAGTTACCAATTCAAATAACATTTGGCTGCTGATTTCGGCTTAATGATCAG 578
QY 347 AAGAAATTTAGCGGAGAGTGTATGATGCTTTTCAAGCTTCTCATTCGACACTTCGTGTTATT 406
DB 579 AAGAACATTAGGCGAAGAGTTTATGATGCTTTAAATGTGCTAATGGCAATGAACATAATT 638
QY 407 GCAAAAGAAAAAGGAGATACCGTGTGATGGCTTTCAAAATACAGATATGAAAAATA 466
DB 639 TCAAGGAAAAAGAAATCAAGTGTGATTGGCTTGGCTTCCCAATTTCTGCTCAGGAATGT 698

QY 467 AAGAAGCTTTGAGGAAGTTCTGTAAGAACTCGTCAACAGATTAGGAACAGAGGCACTC 526
DB 699 CAGATCTCGAGATAGAGAGCAGAGCGCGATAGACGATTAAGCAGAGCGGCCAG 758
QY 527 CTCGAGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACAT 567
DB 759 CTGCAAGAAGCTTCTCTCTACAGCAAAATCGCTTTCAAAAAACCT 799

RESULT 3
US-08-723-415B-5
; Sequence 5, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: Lathangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1110
US-08-723-415B-5

Query Match 7.9%; Score 86.2; DB 2; Length 1157;
Best Local Similarity 56.0%; Pred. No. 7.6e-12;
Matches 191; Conservative 0; Mismatches 138; Indels 12; Gaps 1;

QY 239 GTTTGTGAGAAAGTTGAGCCAAAGGAAGAACATACATACATGAGTTGACAGGAAATT 298
DB 184 GTGTGTGAGAAAGTTGAGCGGAAGGACACAACTTCATACATGAGGTAGCTGATGAGCTG 243
QY 299 TATTCAGAGCTGAAGTCCA-----TGGCACAATATTGGTCAAGGGTTTGTATGAG 346
DB 244 GTATCTGAGTTTACCAACTCAAAATAACCATCTGCAGCTGATTTCGCGAGGCTTATGATCAG 303
QY 347 AAGAAATTTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATTCGACACTTCGTGTTATT 406
DB 304 GAGAACATTAGCAGGAAGAGTTTATGATGCTTTAAATGTACTAATGGCGATGAACATAATT 363

QY 407 GCAAAAGAAAAAGGAGATACGGTGGATGGGCTTCAAAATACAGATATGAAAAATA 466
Db 364 TCAAGGAAAAAGAAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGAATGC 423
QY 467 AGAAGCTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAAGAGGCACTC 526
Db 424 CAGAACTCGAAATCGAGAGCGAGCGGATAGACGGATAAAGCAGAGCGAGCCAG 483
QY 527 CTCAGGAATCGAAAAACAGTTTGATGATCTCCAAAACAT 567
Db 484 CTACAGAATCTCTCTTCAGCAAAATGCTTTTAAAAACCT 524

RESULT 4
US-09-189-627A-5
; Sequence 5, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1110)
US-09-189-627A-5

Query Match 7.9%; Score 86.2; DB 3; Length 1157;
Best Local Similarity 56.0%; Pred. No. 7.6e-12;
Matches 191; Conservative 0; Mismatches 138; Indels 12; Gaps 1;
QY 239 GTTGTGAGAAAGTTGAAGCCAAAGGAAGCAACATACAACTGAGGTTCGACGCAAAAT 298
Db 184 GTGTGTGAGAAAGTTGAGCGGAAGGCAACACTTCATACATGAGGTAGCTGATGAGCTG 243
QY 299 TATTGAGAGCTGAAGTCCA-----TGGCACATATTGGTCAAGGGTTTGATGAG 346
Db 244 GTATCTGAGTTTACCAACTCAATAAACCATCTGGCAGCTGATTTCGCAGGCTTATGATCAG 303
QY 347 AAGAAATATTAGCGGAGAGTGTATGATGCTTCAAGCTTCATTGCACTTCGTGTTATT 406
Db 304 GAGAACATTAGACGAAGAGTTTATGATGCTTTAAATGTACTAATGCGGATGAACATAAAT 363
QY 407 GCAAAAGAAAAAGGAGATACGGTGGATGGGCTTCAAAATACAGATATGAAAAATA 466
Db 364 TCAAGGAAAAAGAAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGAATGC 423
QY 467 AGAAGCTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAAGAGGCACTC 526
Db 424 CAGAACTCGAAATCGAGAGCGAGCGGATAGACGGATAAAGCAGAGCGAGCCAG 483
QY 527 CTCAGGAATCGAAAAACAGTTTGATGATCTCCAAAACAT 567
Db 484 CTACAGAATCTCTCTTCAGCAAAATGCTTTTAAAAACCT 524

RESULT 5
US-09-710-861-5
; Sequence 5, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 5
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1110)
US-09-710-861-5

APPLICANT: La Thangue, Nicholas
APPLICANT: de la Luna, Susana
TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
FILE REFERENCE: 620-54
CURRENT APPLICATION NUMBER: US/09/710,861
CURRENT FILING DATE: 2000-11-13
PRIOR APPLICATION NUMBER: US/09/189,627
PRIOR FILING DATE: 1998-11-10
PRIOR APPLICATION NUMBER: 08/723,415
PRIOR FILING DATE: 1996-09-30
PRIOR APPLICATION NUMBER: GB 9610195
PRIOR FILING DATE: 1996-05-15
NUMBER OF SEQ ID NOS: 25
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 5
LENGTH: 1157
TYPE: DNA
ORGANISM: mouse
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(1110)
US-09-710-861-5
Query Match 7.9%; Score 86.2; DB 3; Length 1157;
Best Local Similarity 56.0%; Pred. No. 7.6e-12;
Matches 191; Conservative 0; Mismatches 138; Indels 12; Gaps 1;
QY 239 GTTGTGAGAAAGTTGAAGCCAAAGGAAGCAACATACAACTGAGGTTCGACGCAAAAT 298
Db 184 GTGTGTGAGAAAGTTGAGCGGAAGGCAACACTTCATACATGAGGTAGCTGATGAGCTG 243
QY 299 TATTGAGAGCTGAAGTCCA-----TGGCACATATTGGTCAAGGGTTTGATGAG 346
Db 244 GTATCTGAGTTTACCAACTCAATAAACCATCTGGCAGCTGATTTCGCAGGCTTATGATCAG 303
QY 347 AAGAAATATTAGCGGAGAGTGTATGATGCTTCAAGCTTCATTGCACTTCGTGTTATT 406
Db 304 GAGAACATTAGACGAAGAGTTTATGATGCTTTAAATGTACTAATGCGGATGAACATAAAT 363
QY 407 GCAAAAGAAAAAGGAGATACGGTGGATGGGCTTCAAAATACAGATATGAAAAATA 466
Db 364 TCAAGGAAAAAGAAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGAATGC 423
QY 467 AGAAGCTTGAGGAAGTTCGTAAGAACTCGTCAACAGATTAGGAACAAAGAGGCACTC 526
Db 424 CAGAACTCGAAATCGAGAGCGAGCGGATAGACGGATAAAGCAGAGCGAGCCAG 483
QY 527 CTCAGGAATCGAAAAACAGTTTGATGATCTCCAAAACAT 567
Db 484 CTACAGAATCTCTCTTCAGCAAAATGCTTTTAAAAACCT 524
RESULT 6
US-09-640-211A-566
; Sequence 566, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 566
; LENGTH: 358
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-566

Query Match 7.8%; Score 85; DB 4; Length 358;
Best Local Similarity 62.5%; Pred. No. 9.3e-12;
Matches 159; Conservative 0; Mismatches 80; Indels 15; Gaps 1;

QY 209 GGTGGGGGTCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTGAAGCAAGGAAGA 268
DB GGAGAGGACTACGCATTCAGCATGAAGTTTGAAGAAAGTGAGAGCAAGGGCTGG 117

QY 269 ACAACATACATAGGTTCAGACGAAATTT-----ATTGAGAGCTGAAG 313
DB 118 ACAACATACAGAGGTTCATCTGAATTAGTGGCCGAATTTGTGAATCCAAACAGACA 177

QY 314 TCATGGCCACATATTGGTCAAGGTTTGTGAGAGAAATATTAGGCGGAGAGTGTATGAT 373
DB 178 CATCTTTCACAGATCAGCAACAGTTTGTGAGAGAAACATTTAGGAGGAGGGTGTATGAT 237

QY 374 GCTTTCAACGTTCTCATTTGACATTCGTTTATTGCAAAAGAAAAGAGAGATACGGTGG 433
DB 238 GCACTGACGTACTGATGGCATGGACATATATCAAGAGAAAAGAGATAGATCG 297

QY 434 ATGGGCTTTCAA 446
DB 298 AAGGGCTACCTA 310

RESULT 7
US-08-428-131-1
; Sequence 1, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4100
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 55..1284
; US-08-428-131-1

Query Match 7.6%; Score 83.2; DB 2; Length 1700;
Best Local Similarity 53.7%; Pred. No. 5.1e-11;
Matches 203; Conservative 0; Mismatches 163; Indels 12; Gaps 1;

QY 202 GATCACCGGTTGGGGCTCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTCGAAGCAA 261
DB 384 GAAGAATGGCAAGGGCTCGGCAATTTCTCCATGAAGTGTGTGAGAAAGTTCAGAGGAA 443

QY 262 AGGAAGAACAAACATACAAATGAGGTTGCAGACGAAATTTATTTCAGAGCTGAAGTC----- 315
DB 444 AGGAACCACTCTCTACAATGAGGTGGTGCAGAGTGGTGGCAGAGTTTCAGCGCTGCCGA 503

QY 316 -----CATGGCACATATTGGTCAAGGTTTGTATGAGAAATATTAGGCGGAGAGTGA 369
DB 504 CAACCAATCTTCAACAAACGAATCAGCTTATGACCAAGAAACATCCGGCGGGTGTCTA 563

QY 370 TGATGCTTTCAACGTTCTCATTCGCTCTGTTTATTGCAAAAGAAAAGAGAGATACG 429
DB 564 CGATGCTTTAAATGTGCTAATGCCATGAACATCATCTCCAAGAGAAAGAGAGATCAA 623

QY 430 GTGGATGGGCTTTCAAATTTACAGATATGAAAAAATAAGAAAGCTTTGAGGAAGTTCTGTA 489
DB 624 ATGATCGGCTCTCCACCACCACTCAGCTCAGGAGTCCAGAACTTAGAGGTGGAGAGGCA 683

QY 490 AGAATCGTCAACAGATTAGGAACAAGAGGCACTCTCCAGAAATCCAAAAACAGTT 549
DB 684 GAGGAGGCTGGAGAGGATCAACAGAGCAGTCTCAGCTCCAGGAGCTCATCTCTGCAGCA 743

QY 550 TGATGATCTCCAAAACAT 567
DB 744 AATTGCTTCAAGAACTT 761

RESULT 8
US-08-602-846-1
; Sequence 1, Application US/08602846
; Patent No. 5871901
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas B
; TITLE OF INVENTION: ASSAY FOR INHIBITORS OF DP-1 AND OTHER DP
; NUMBER OF SEQUENCES: 3
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye PC
; STREET: 8th Floor, 1100 No. 5871901th Glebe Road
; CITY: Arlington
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/602,846
; FILING DATE: 26-FEB-1996
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: CRAWFORD, ARTHUR R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 620-12
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1700 base pairs
; TYPE: nucleic acid and encoded amino acids
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cdna
; US-08-602-846-1

Query Match 7.6%; Score 83.2; DB 2; Length 1700;
Best Local Similarity 53.7%; Pred. No. 5.1e-11;
Matches 203; Conservative 0; Mismatches 163; Indels 12; Gaps 1;

QY 202 GATCACCCTGGGGCTCCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAA 261
|||
Db 384 GAAGATGGCAAGGGCTCGGCATTCTCCATGAAGGTGTGTGAGAGGTGCAGAGGAA 443
|||
QY 262 AGAAGAACACATACAAATGAGGTTGCAGACGAAATTTATTCAGAGCTGAAGTC----- 315
|||
Db 444 AGAACCACCTCTACAAATGAGGTGCTGACGAGCTGGTGGCAGAGTTCAAGCGCTGCCGA 503
|||
QY 316 -----CATGGCACATATTGGTCAAGGTTTGTATGAGAGAAATATTAGCGGAGAGTGA 369
|||
Db 504 CAACACATTTCTACCAACGAATCAGCTTAAGCCAGAGAAATCCGGCGGGGTGTCTA 563
|||
QY 370 TGATGCTTTCAAGCTTCTCATTTGCACCTTCGTGTTATTGCAAAAGAAAAAGGAGATACG 429
|||
Db 564 CGATGCTTTAAATGTCTAATGCGCATGACATCATCTCCAGGAGAGAGAGATCAA 623
|||
QY 430 GTGGATGGGCTTTCAAATTTACAGATATGAAAAAATAAGAAAGCTTGAGGAAGTTGCTAA 489
|||
Db 624 ATGGATCGGCTTCCCAACCAACTCAGCTCAGGAGTGCAGAACTTAGAGGTGGAGAGGCA 683
|||
QY 490 AGAATCGTCAACAGATTAGGACAGAGGAGGCACTCTCCAGGAATCGAAAAACAGTT 549
|||
Db 684 GAGGAGGCTGGAGAGGATCAACAGAGGAGTCTCAGCTCCAGGAGCTATCTCTGCAGCA 743
|||
QY 550 TGATGATCTCCAAACAT 567
|||
Db 744 AATTGCTTCAGAACTT 761
|||
RESULT 9
US-09-078-596-1
; Sequence 1, Application US/09078596
; Patent No. 6150116
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 6150116th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/078,596
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1700 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS

LOCATION: 55..1284
US-09-078-596-1
Query Match 7.6%; Score 83.2; DB 3; Length 1700;
Best Local Similarity 53.7%; Pred. No. 5.1e-11;
Matches 203; Conservative 0; Mismatches 163; Indels 12; Gaps 1;
QY 202 GATCACCCTGGGGCTCCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAA 261
|||
Db 384 GAAGATGGCAAGGGCTCGGCATTCTCCATGAAGGTGTGTGAGAGGTGCAGAGGAA 443
|||
QY 262 AGAAGAACACATACAAATGAGGTTGCAGACGAAATTTATTCAGAGCTGAAGTC----- 315
|||
Db 444 AGAACCACCTCTACAAATGAGGTGCTGACGAGCTGGTGGCAGAGTTCAAGCGCTGCCGA 503
|||
QY 316 -----CATGGCACATATTGGTCAAGGTTTGTATGAGAGAAATATTAGCGGAGAGTGA 369
|||
Db 504 CAACACATTTCTACCAACGAATCAGCTTAAGCCAGAGAAATCCGGCGGGGTGTCTA 563
|||
QY 370 TGATGCTTTCAAGCTTCTCATTTGCACCTTCGTGTTATTGCAAAAGAAAAAGGAGATACG 429
|||
Db 564 CGATGCTTTAAATGTCTAATGCGCATGACATCATCTCCAGGAGAGAGAGATCAA 623
|||
QY 430 GTGGATGGGCTTTCAAATTTACAGATATGAAAAAATAAGAAAGCTTGAGGAAGTTGCTAA 489
|||
Db 624 ATGGATCGGCTTCCCAACCAACTCAGCTCAGGAGTGCAGAACTTAGAGGTGGAGAGGCA 683
|||
QY 490 AGAATCGTCAACAGATTAGGACAGAGGAGGCACTCTCCAGGAATCGAAAAACAGTT 549
|||
Db 684 GAGGAGGCTGGAGAGGATCAACAGAGGAGTCTCAGCTCCAGGAGCTATCTCTGCAGCA 743
|||
QY 550 TGATGATCTCCAAACAT 567
|||
Db 744 AATTGCTTCAGAACTT 761
|||
RESULT 10
US-09-640-211A-465
; Sequence 465, Application US/09640211A
; Patent No. 6833446
; GENERAL INFORMATION:
; APPLICANT: Wood, Marion
; APPLICANT: Shenk, Michael A.
; APPLICANT: McGrath, Annette
; APPLICANT: Glenn, Matthew
; TITLE OF INVENTION: Compositions and Methods for the
; TITLE OF INVENTION: Modification of Gene Transcription
; FILE REFERENCE: 11000.1021C1U
; CURRENT APPLICATION NUMBER: US/09/640,211A
; CURRENT FILING DATE: 2000-08-16
; NUMBER OF SEQ ID NOS: 2368
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 465
; LENGTH: 364
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-640-211A-465
Query Match 7.6%; Score 82.8; DB 4; Length 364;
Best Local Similarity 57.9%; Pred. No. 3.4e-11;
Matches 175; Conservative 0; Mismatches 112; Indels 15; Gaps 1;
QY 209 GGTGGGGCTCCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCCAAAGGAGA 268
|||
Db 57 GGGAGGGGACTTCGCCAGTTTCAGCATGAAAGTATGTCAAAGGTTCAGAGCAAGGGTCGG 116
|||
QY 269 ACAACATACATGAGGTTTCAGACGAAATTTATTTCAGAGCTGAA-----G 313
|||
Db 117 ACCAGTATATGAGGTTTCAGATGAAATTTAGTTTCAGATATATCAATCTTAACAGTGG 176
|||
QY 314 TCCATGGCACATATTGGTCAAGGTTTGTATGAGAGAAATATTAGCGGAGAGTGTATGAT 373
|||
Db 177 CTCAATTTCTCCTGATCAGCAACAAATATGATGAGAAAAACATAGGGGGAGGGTGTATGAT 236
|||

QY 374 GCTTTCAAGCTTCTCATTGCGACTTCTGTTATTGCAAAAGAAAAAGGAGATACGGTGG 433
Db 237 GCATTGAATGTACTGATGGCAATGGACATCATATCAAGGACAAAGAAATTCAGTGG 296
QY 434 ATGGGCTTTCAAAATTACAGATATGAAAAATAAAGAACTTCAGGAAGTTCCGTAAGAA 493
Db 297 AAGGGTTACCTAGCAAGTCCATATGACCTTGAAGACTTGAAGCAAGCCGATGGGA 356
QY 494 CT 495
Db 357 TT 358

RESULT 11
US-08-723-415B-3
; Sequence 3, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: Lathangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Giebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1154 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1107
US-08-723-415B-3

Query Match 7.5%; Score 81.2; DB 2; Length 1154;
Best Local Similarity 55.0%; Pred. No. 1.4e-10;
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;
QY 239 GTTTGTGAGAAAGTTGAAGCAAGGAAGAAACAACATACAAATCAGGTTGCGAGCAAAATT 298
Db 184 GTGTGTGAGAAAGTTGAGGAAAGGCAACAACTTCATACAAATGAGTGTGATGAGCTG 243
QY 299 TATTCAGAGCTGAAGTCCATGGCACAATTT-----GGTCAAGGGTTTGTATGAGAAG 349
Db 244 GTATCTGAGTTTACCAACTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAG 303

QY 350 AATATTAGCGGAGAGTGTATGATGCTTCAAGGTTCTCATTGCACCTTCGTGTTATTGCA 409
Db 304 AACATTAGACGAAGAGTTTTATGATGCTTTAAATGTACTAATATGCGGATGAACATAATTTC 363
QY 410 AAAGAAAAAAGGAGATACGGTGGATGGCGCTTTCAAATTTACAGATATGAAAAATAAAG 469
Db 364 AAGAAAAAAGAAATCAAGTGGATTGGCTGCTCCCAATTTCTGCTCAGGATGCCAG 423
QY 470 AAGCTTTGAGGAAGTTTCGTAAAGAACTTCGTCAACAGATTAGGAACAAGAGGCACTCTC 529
Db 424 AACCTGGAATCGAGAAAGCAGAGCGGATAGAACGGATAAAGCAGAAAGCGAGGCCAGCTA 483
QY 530 CAGGAATCGAAAAACAGTTTGATGATCTCCAAAAACAT 567
Db 484 CAAGAACTTCTCTCTTCAGCAAAATTGCTTTTAAAAAACCT 521

RESULT 12
US-09-189-627A-3
; Sequence 3, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1154
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1107)
US-09-189-627A-3

Query Match 7.5%; Score 81.2; DB 3; Length 1154;
Best Local Similarity 55.0%; Pred. No. 1.4e-10;
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;
QY 239 GTTTGTGAGAAAGTTGAAGCAAGGAAGAAACAACATACAAATCAGGTTGCGAGCAAAATT 298
Db 184 GTGTGTGAGAAAGTTGAGGAAAGGCAACAACTTCATACAAATGAGTGTGATGAGCTG 243
QY 299 TATTCAGAGCTGAAGTCCATGGCACAATTT-----GGTCAAGGGTTTGTATGAGAAG 349
Db 244 GTATCTGAGTTTACCAACTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAG 303
QY 350 AATATTAGCGGAGAGTGTATGATGCTTCAAGGTTCTCATTGCACCTTCGTGTTATTGCA 409
Db 304 AACATTAGACGAAGAGTTTTATGATGCTTTAAATGTACTAATATGCGGATGAACATAATTTC 363
QY 410 AAAGAAAAAAGGAGATACGGTGGATGGCGCTTTCAAATTTACAGATATGAAAAATAAAG 469
Db 364 AAGAAAAAAGAAATCAAGTGGATTGGCTGCTCCCAATTTCTGCTCAGGATGCCAG 423
QY 470 AAGCTTTGAGGAAGTTTCGTAAAGAACTTCGTCAACAGATTAGGAACAAGAGGCACTCTC 529
Db 424 AACCTGGAATCGAGAAAGCAGAGCGGATAGAACGGATAAAGCAGAAAGCGAGGCCAGCTA 483
QY 530 CAGGAATCGAAAAACAGTTTGATGATCTCCAAAAACAT 567
Db 484 CAAGAACTTCTCTCTTCAGCAAAATTGCTTTTAAAAAACCT 521

RESULT 13

US-09-710-861-3
; Sequence 3, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; PRIORITY FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1154
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1107)
US-09-710-861-3

Query Match 7.5%; Score 81.2; DB 3; Length 1154;
Best Local Similarity 55.0%; Pred. No. 1.4e-10;
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

QY	239	GTGTTGAGAAAGTTGAGCCAAAGGAAGCAACATACATGAGGTTCGAGACGAAATT	298
DB	184	GTGTTGAGAAAGTTGAGCCAAAGGAAGCAACATTCATGAGGTTCGAGACGAAATT	243
QY	299	TATTCAGAGCTGAAGTCCATGCGACATATT-----GGTCAAGGGTTTGTATGAGAAG	349
DB	244	GTATCTGAGTTTACCAACTCAATATACCATCTGGCAGCTGATTCGCTTATGATCAGGAG	303
QY	350	ATATTTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATGCACTTCGTGTTATGCA	409
DB	304	ATCATTTAGACGAGAGTTTATGATGCTTTTAAATGTACTAATGGCGATGAACATAATTCA	363
QY	410	AAAGAAAAAAGAGATACGGTGGATGGGCTTTTCAAAATTTACAGATATGAAAAAATAAG	469
DB	364	AGGAAAAAAGAAATCAAGTGATGGCTGCTTACCAATTTCTCTCAGGAATGCCAG	423
QY	470	AAGCTTTAGGAAAGTTCTGTAAGAACTCGTCAAGATTTAGGAACAAAGAGGCACTCTTC	529
DB	424	AACCTGGAATTCGAGAAGCAGAGCGGATAGAACGGATAGAAAGCAGAGCGGCCAGCTA	483
QY	530	CAGGAATCGAAAAACAGTTTGTATGATCTCCAAACAT	567
DB	484	CAAGAACTTCTCTTCAGCAAAATGCTTTTAAAAACCT	521

RESULT 14

US-08-723-415B-7
; Sequence 7, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: LaThangue, Nicholas B.
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA

ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1202 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1155
US-08-723-415B-7

Query Match 7.5%; Score 81.2; DB 2; Length 1202;
Best Local Similarity 55.0%; Pred. No. 1.4e-10;
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

QY	239	GTGTTGAGAAAGTTGAGCCAAAGGAAGCAACATACATGAGGTTCGAGACGAAATT	298
DB	232	GTGTTGAGAAAGTTGAGCCAAAGGAAGCAACATTCATACATGAGGTTCGAGACGAAATT	291
QY	299	TATTCAGAGCTGAAGTCCATGCGACATATT-----GGTCAAGGGTTTGTATGAGAAG	349
DB	292	GTATCTGAGTTTACCAACTCAATATACCATCTGGCAGCTGATTCGCTTATGATCAGGAG	351
QY	350	ATATTTAGCGGAGAGTGTATGATGCTTTCAAGCTTCTCATGCACTTCGTGTTATGCA	409
DB	352	ATCATTTAGACGAGAGTTTATGATGCTTTTAAATGTACTAATGGCGATGAACATAATTCA	411
QY	410	AAAGAAAAAAGAGATACGGTGGATGGGCTTTTCAAAATTTACAGATATGAAAAAATAAG	469
DB	412	AGGAAAAAAGAAATCAAGTGATGGCTGCTTACCAATTTCTCTCAGGAATGCCAG	471
QY	470	AAGCTTTAGGAAAGTTCTGTAAGAACTCGTCAAGATTTAGGAACAAAGAGGCACTCTTC	529
DB	472	AACCTGGAATTCGAGAAGCAGAGCGGATAGAACGGATAGAAAGCAGAGCGGCCAGCTA	531
QY	530	CAGGAATCGAAAAACAGTTTGTATGATCTCCAAACAT	567
DB	532	CAAGAACTTCTCTTCAGCAAAATGCTTTTAAAAACCT	569

RESULT 15

US-09-189-627A-7
; Sequence 7, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415

```
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1202
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1155)
US-09-189-627A-7

Query Match      7.5%; Score 81.2; DB 3; Length 1202;
Best Local Similarity 55.0%; Pred. No. 1.4e-10;
Matches 186; Conservative 0; Mismatches 143; Indels 9; Gaps 1;

QY 239 GTTGTGAGAAAGTTGAAGCCAAAGGAAGCAACATACATGAGGTTGCAGACGAAATT 298
Db 232 GTGTGTGAGAAAGTTGAGCGGAAAGGCACAACTTCATACATGAGGTAGCTGATGAGCTG 291
QY 299 TATTTCAGAGCTGAAGTCCATGGCACATATT-----GGTCAAGGGTTTGATGAGAAG 349
Db 292 GTATCTGAGTTTACCACCTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAG 351
QY 350 AATATTAGCGGAGAGTGATGATGTTTCAACGTTCTCATTTGCACTTCGTGTTATTGCA 409
Db 352 AACATTAGACGAAGAGTTTATGATGCTTTAAATGTACTAATGGCGATGAACATAATTTC 411
QY 410 AAAGAAAAGAGAGATACGGTGGATGGCGCTTTCAAATTACAGATATGAATAAATAAG 469
Db 412 AAGGAAAAAAGAAATCAAGTGGATGGCGCTGCCCTACCAATTCTGCTCAGGAATGCCAG 471
QY 470 AAGCTTGAGGAAGTTCTGTAAGAACTCGTCAACAAGATTAGGAACAAGAGGCACCTCCTC 529
Db 472 AACCTGGAATCGAAGCAGAGCGCGGATAGACGGTAAAGCAGAGCGAGGCCAGCTA 531
QY 530 CAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAAT 567
Db 532 CAAGAACTTCTCCTTCAGCAAAATTGCTTTTAAAAAAGCT 569
```

Search completed: February 6, 2005, 13:09:14
Job time : 230 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: February 6, 2005, 13:09:21 ; Search time 690 Seconds
(without alignments)
9084.062 Million cell updates/sec

Title: US-10-088-830-1
Perfect score: 1089
Sequence: 1 gaattcgacgacgcgcaaa.....aaaaaaaaaaaaaaaaaaaaa 1089

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA.*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
- 6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq.*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq.*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq.*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq.*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	452.6	41.6	1406	18	US-10-437-963-33888 Sequence 33888, A
2	385.4	35.4	774	18	US-10-425-115-4115 Sequence 4115, Ap
3	203.4	18.7	476	18	US-10-767-701-17642 Sequence 17642, A
4	150.8	13.8	1639	18	US-10-739-930-1190 Sequence 1190, Ap
5	141.6	13.0	1431	17	US-10-424-599-43806 Sequence 43806, A
6	139.8	12.8	1585	18	US-10-425-115-87351 Sequence 87351, A
7	133.4	12.2	1483	18	US-10-425-115-2033 Sequence 2033, Ap
8	131.8	12.1	1350	17	US-10-425-114-4049 Sequence 4049, Ap
9	124.6	11.7	1167	17	US-10-425-114-20051 Sequence 20051, A
10	121.4	11.1	1751	17	US-10-424-599-43105 Sequence 43105, A
11	120	11.0	911	17	US-10-425-114-36537 Sequence 36537, A

12	102.8	9.4	642	18	US-10-767-795-6106 Sequence 6106, Ap
13	99.2	9.1	2968	14	US-10-198-846-10136 Sequence 10136, A
14	98.8	9.1	2320	17	US-10-641-643-1456 Sequence 1456, Ap
15	97	8.9	1266	9	US-09-954-456-1176 Sequence 1176, Ap
16	97	8.9	2968	18	US-10-357-930-20983 Sequence 20983, A
17	97	8.9	2968	18	US-10-357-930-26827 Sequence 26827, A
18	91.2	8.4	1213	18	US-10-437-963-63675 Sequence 63675, A
19	88	8.1	1832	9	US-09-220-091-8 Sequence 8, Appli
20	85.4	7.8	792	18	US-10-437-963-64593 Sequence 64593, A
21	85	7.8	358	18	US-10-856-499-566 Sequence 566, App
22	82.8	7.6	364	18	US-10-856-499-465 Sequence 465, App
23	76.6	7.0	738	17	US-10-424-599-91931 Sequence 91931, A
24	74.8	6.9	692	15	US-10-206-901B-21 Sequence 21, Appl
25	74.8	6.9	1440	17	US-10-172-118-1357 Sequence 1357, Ap
26	74.8	6.9	1440	17	US-10-160-554-4 Sequence 4, Appli
27	74.8	6.9	1440	17	US-10-342-887-1357 Sequence 1357, Ap
28	74.8	6.9	1440	17	US-10-641-643-1399 Sequence 1399, Ap
29	74.8	6.9	1440	18	US-10-717-597-73 Sequence 73, Appl
30	74.8	6.9	2607	15	US-10-106-698-569 Sequence 569, App
C 31	71.6	6.6	397	9	US-09-960-352-13784 Sequence 13784, A
C 32	69.6	6.4	1571	17	US-10-424-599-11317 Sequence 11317, A
C 33	69.4	6.4	424	9	US-09-960-352-11218 Sequence 11218, A
C 34	69.4	6.4	837	18	US-10-425-115-178785 Sequence 178785, A
C 35	69.4	6.4	1461	9	US-09-811-367B-6 Sequence 6, Appli
C 36	69.2	6.4	780	18	US-10-425-115-55942 Sequence 55942, A
37	69	6.3	547	18	US-10-021-323-8671 Sequence 8671, Ap
38	68.8	6.3	1649	18	US-10-821-273-19 Sequence 19, Appl
39	68	6.2	3513	17	US-10-374-979-69 Sequence 69, Appl
40	68	6.2	3513	17	US-10-182-936A-69 Sequence 69, Appl
41	68	6.2	3513	18	US-10-477-238A-648 Sequence 648, App
42	68	6.2	3513	18	US-10-680-287A-648 Sequence 648, App
43	68	6.2	3513	18	US-10-723-860-293 Sequence 293, App
44	68	6.2	4649	18	US-10-723-860-5013 Sequence 5013, Ap
45	67.8	6.2	1113	18	US-10-425-115-118211 Sequence 118211, A

ALIGNMENTS

RESULT 1

US-10-437-963-33888/c
; Sequence 33888, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33888
; LENGTH: 1406
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37957C.1
US-10-437-963-33888

Query Match 41.6%; Score 452.6; DB 18; Length 1406;
Best Local Similarity 74.7%; Pred. No. 3.1e-98;
Matches 639; Conservative 0; Mismatches 199; Indels 17; Gaps 5;
QY 248 AAAGTTGAAGCCAAAGGAAGAACATCAATAGAGTTGCAGCAAAATTTATTCAGAG 307
Db 854 AACGATGAAGTCTATACAAAGCTCAGGAGAGCTTATAGTTGCCGATGATTTTTCGGAG 795

308 CTGAAGTCCATGCGACATATTTGCTCAAGGGTTTGTAGAGAAATATTTAGGCGGAGAGTG 367
Db |||||
794 CTGAAGTCCATTAACGAGAACGGTCTCGAGTTTGTAGAGAAATATTTAGGCGGAGAGTA 735
QY |||||
368 TATGATGCTTTCAAGCTTCTCATTGACATTCGTTGTTATTCGAAAGAAAGAAAGAGAGATA 427
Db |||||
734 TATGATGCTTTCAAGTGTCTCTGCAATTCGTTGTTATTCGAAAGAAATTAAGAGAGATA 675
QY |||||
428 CGTGGATGGCCCTTTCAATTAACAGATATGAAATTAAGAAAGCTTGAAGAAAGTTTCGT 487
Db |||||
674 AAGTGGATGGCCCTTACTAATATAGATACGAAAGATACAGAGTTGGAGAAAGTTTCCAC 615
QY |||||
488 AAGAACTCTGTAACAGATTAAGAAACAGAAAGGCACTCTCCAGGAAATTCGAAAGAAACAG 547
Db |||||
614 AAGAACTCTATCAACAGGATCAAGAAATTAAGAAAGCTTCTCCAGGAAATTTGAAGAGCAG 555
QY |||||
548 TTTGATGATCTCCAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAGAAAT 607
Db |||||
554 TTTGATGATCTTCAAGATATTAATTAAGCAACAGGCTAGTCAGAGGCGCCAGCAGAAAGT 495
QY |||||
608 GTTAATGGCATCGCCCTTCATTCGTAATGTTGTTCAAGACATCTAGGAAAGCAAGGGTGGAA 667
Db |||||
494 GTTAATGGCATCTCTCTCCGTTCTTATTGATCAAGACATCTCCGAAAGCAAGGGTGGAA 435
QY |||||
668 ATTGAGATTCAGATGACTCGAAGTTTGCCCATTTGAGTTCAATGGTGCACCAATTCACA 727
Db |||||
434 ATTGAGATTCGGAAGATTCGAAGTTTGCAACGGTTTCGACTTCAACGGTGCACCAATTCACC 375
QY |||||
728 TTGATGATGATCTCTCAATCTCTGAGGGGTAAAGCGTAAACAGCATAGGAAGAGCTGGC 787
Db |||||
374 ATGATGATGATGATGATCAATCTTTGAAGCCATCAGGCGTTAAACAAAGGAAGAGCTGGC 315
QY |||||
788 CGCGCCACCCCTTCACT-AGAGACTCAAGAAATATTACAAA-----TGAATTAAGAGTGTAG 842
Db |||||
314 CTCTCCATTCACCTTAAGAGGCGACAAAGATGTGACAAACCAATTGAAGTGTCAAACTCAA 255
QY |||||
843 AACTGCACACCGGATCTTTTGGCAGCTATGTATAGCTATA-----TATCTCAT 895
Db |||||
254 AACTGGCACACCACTTTCTTTTGGCAGCTATGTATAGCTATAGCTATACCTTTTCAGTAT 195
QY |||||
896 GAAACTTGACCTAGTTTATAGACAGCTCTCTCAGGCTTGAGAAGA-TTTTAACTTCGCAA 954
Db |||||
194 GGAACCTCGACCTAGTTTATAGACAGCTCTCTCAGGCTTTTATAGCAGATATGACCTTGGCA 135
QY |||||
955 ATTTGCTCTCTTTTGTGCTAGCAGGTTATTTAGTCTCAGATAGATGATTCATATATG 1014
Db |||||
134 ATTTGCTCTCTTTTGTGCTAGCAGGTTGA-----GTCCTGGAATAGTTTCTTCTGCGTG 79
QY |||||
1015 TGTGCTATGAAACATTTGATAGCAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1074
Db |||||
78 TGTGTTATGTATAAACTCAATTCGCTTTCAATTCGAAGCTTGTGTACAAAGCTTCCATTGA 19
QY |||||
1075 AAAAAAAGAAAAA 1089
Db |||||
18 TGAAGAAAAAAGAAAA 4

RESULT 2

US-10-425-115-4115
; Sequence 4115, Application US/10425115
; Publication No. US2004021427A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants
; FILE REFERENCE: 38-21(5322)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326

; SEQ ID NO 4115

; LENGTH: 774

; TYPE: DNA

; ORGANISM: Zea mays

; FEATURE:

; OTHER INFORMATION: Clone ID: MPT4577_103754C.1

US-10-425-115-4115

Query Match 35.4%; Score 385.4; DB 18; Length 774;

Best Local Similarity 74.0%; Pred. No. 3.2e-82;

Matches 547; Conservative 0; Mismatches 176; Indels 16; Gaps 4;

QY 53 GCTACCGCGCTGACCTGACCGCGGTGACATTTCTCGAAGCTTCCAGTGTCCCGCCG 112

Db 18 GCTGCGCGCGACCGCGGTGACCTGACCGCTGCGCATCAGCGAAGAGCTGGGATCCCTCCG 77

QY 113 CTTCCCGAA-----GCGGCGGTAATGCGTCCAAAGGAAGGGGCTGTTGACCGG 163

Db 78 CTTCCCGGACGGGGAAGCGTCCGCGCAATGCACTGAGGGAGAGAGGGGTGATAAG 137

QY 164 GATAAGATAGGAAGAGGAGGCTCGCGGCAACGAGGATCACCGGTTGGGGGCTCCG 223

Db 138 AACAGAAAGCGGAGGAGGTAAGGCTGCTACGACGCGATCGCGGCTGGGGCTCCG 197

QY 224 GAGTACAGCAATAGTTTGTGAGAAAGTTGAGCCAAAGGAAGAACACATCAATGAG 283

Db 198 GAGTTCAGCAAGATAGTTTCTAAAGGTTGAGACCAGGCGGACTACTTACAGTGAG 257

QY 284 GTTGACAGCAATTTTATTCAGAGCTGAAAGTCCATGGCAGATATTTGCTCAAGGGTTGAT 343

Db 258 GTTGACAGTGAATTTGTGAGAGCTGAAGT---TGACTCTTATTTGTTCAAGAGTTGAT 314

QY 344 GAGAAATATTAGGCGGAGGTGATGATGCTTTCAAGCTTCTCAATTCACATGATGATG 403

Db 315 GAGAAATATCAGGAGGAGGTGATGATGCTTTTAAATGCTAAATTTGCACTACGTTG 374

QY 404 ATTGCAAAAGAAAGAGGATACGTTGGTGGGCTTTCAATTTACAGATATGAAAA 463

Db 375 ATCAAAAGATGAAAGAGATAAATGGATGGGCTTGTCTAATTTCCAATATGAAAG 434

QY 464 ATAAAGAGCTTGAGGAGTTCGTAAGAACTCGTCAACAGATTAAGGAACAGAAAGGCA 523

Db 435 AT---TCAGTTGGAGGAACTCGAAAGAACTCATGATCAGGATTAAGAACAGAAAAA 491

QY 524 CTCCTCCAGAAATCGAAA-AACAGTTTATGATCTCCTCAAAACATCAAGTTAGTAAACA 582

Db 492 CTTCTCCAGAAATTTGAAAGACCCAGCGTATGATGACCTCCAGAAATATCAAGTTTCGCAACA 551

QY 583 AACACTGGAAGCTCAGCAGAGAAATGTTAATGGCATCCGCTTCCATTCGTTATTTGTTCAA 642

Db 552 GTTTCTACAGGCGCAGCAGAGTGTGAATGGTATCTGCTTCCATTTCTTATTTGTTCAA 611

QY 643 GACATCTAGGAAAGCAAGGGTGGAAATTTGAGATTTTCAGATGACTCGAAAGTTTCCCATTT 702

Db 612 GGCATCCAGAAAGCAAGGGTGGAGATTGAGATCTCAGAAAACTCAAGAGTTTCTGTTGTT 671

QY 703 CGAGTTCAATGGTGCACCATTTCACTTGCATGATGATCTCTCAATCTTGGAGGGGTAAG 762

Db 672 GGACTTCAACTGCTACCATTTCACTTGCACGAGATGTATCAATCTTTGAAGCCATCAG 731

RESULT 3

US-10-767-701-17642
; Sequence 17642, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei

QY	272	ACATACAATCAGGTTGCAGACGAAATTTATTTCAGAGCTG-----AAGTCC	316
DB	365	ACATACAATCAGGTTGGCGGATGAGCTTGTCTGAATTTCTGAACCAACGAATAGTGAA	424
QY	317	ATGCGCATATTTGGTCAAGGGTTTGATGAGAAGAAATATTAGCGCGGAGAGTGTATGATGCT	376
DB	425	TTGGCCCTCTGATCAGCAACAATATGATGAAAAAACATCCGCCGAAGGGTCTATGATGCT	484
QY	377	TTGAAAGCTTCTCATTTGCACCTTCGGTTGTTATTCGAAAGAAAAAGGAGATACGGTGGATG	436
DB	485	CTGAACGTTCTCATGGCAATGGATATTATTTCGAAGGACAAAAAGGAAATTCATGSGGG	544
QY	437	GGCCTTTCAAAATTACAGATATGAAAAAATAAAGAAGCTTGAGGAAGTTTCGTAAGAATCT	496
DB	545	GGCCTTCTCGTACTCTGTGAATGATATTGAAGAACTTAAAGACAGAGCGCTTTGGGCTC	604
QY	497	GTCAACAAGATTAGGAAACAAGAGGCACCTCTCCAGGNAATCGA-AAAAACAGTTTGATGA	555
DB	605	AGGAATAGAATTGAAGAAAAAGAAACAGCCCTATCTGCAGGAGCTTGAGGGAGCAATTCATAGG	664
QY	556	TCTCCAAAAACATCAAGTTTACGTACCAACCAACTCGAAAGCTCAGCAGAGAAATGTTAATGG	615
DB	665	TCTTCAGAACCTTATTCAACGAACAGCAGCTTATATAGCTCAAGAAATCTCCCAAGTG	724
QY	616	CA--TCGCGCCTTCCATTCGTAATTTGGTCAAGACATCTAGGAAGCAGGGTGGAAATTGAG	673
DB	725	AAGTGATCTCTTACCCCTTATTTTGGTACAGACACGCCCTCATGCAACTCGGGGGAATGAA	784
QY	674	ATTTTCAGATGACTCGAAGTTTGGCCATTTTCGAGTTCAATGGTGCACCACTTACATTCGAT	733
DB	785	ATAACAGAAGATATGCACCTTGTTCATTTTGAATTCATAGCACTCCCTTTTGAGCTGCAT	844
QY	734	GATGATCTCTCAATCCCTTGAGG	755
DB	845	GACGACAAATTATGTTCTCAAGG	866

```

RESULT 6
US-10-425-115-87351
; Sequence 87351, Application US/10425115
; Publication NO. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 87351
; LENGTH: 1585
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRR4577_179669C.1
US-10-425-115-87351

```

QY	332	----	CAAGGGTTT	TGATGAGAAGAA	TATTAGCGGA	362
Db	254	AGGCAC	CAGATCT	GTATAACCC	TAAACCGGCAACAATATGATGAGAA	313
QY	363	GAGTGTAT	GATGCTTTCAAC	GGTTCTCATTTGC	ACTTCGTGTTATTGCAAAAGAAAAGG	422
Db	314	GAGTTTAT	GATGCTTTGAAT	GTGTTCTGATGGCTAT	GGACATTATATCTAAAGATAAAAAGG	373
QY	423	AGATA	CGGTGGATGGG	CCITTCAAATTACAGATAT	TGAABAAATAAAGAGCTTTGAGGAAG	482
Db	374	AGATCC	AGTGGAAAGGGCTTC	CCGCGTACTAGTATAGTGA	CATTGAGAATGAAGACGG	433
QY	483	TTCGTAA	AGAACTCGTCAACA	GAGATTAGGAAC	AGGAGGCACCTCCTCCAGGAATCGAAA	542
Db	434	AGCTTTGG	GACTCAAAAGGTAGG	ATTGAAAGAAAAGTGCTTACCTACAGGAGCTACAAG	493	
QY	543	AACAGTTT	GATGATCTCCAAA	AAACATCAAGTTACGTAA	CCAAAACACTCGGAAGCTTCAGCAG	602
Db	494	ACCAATAT	GATGGTCTGCTCAAA	AACCTGATTCAAAGAAATGAGCA	TATATGTTTCAGGAA	553
QY	603	A---	GAATGTTAATGGCATCG	CCCTCCATTCGTATTGGTCAAGACA	TCTAGGAAAGCAA	659
Db	554	ACAGCC	TTCTGGTGGAGTGGCTTTG	CCATTATCCTTAGTTTCAGAC	CCGACCTCATGCTA	613
QY	660	GGGTGG	AAAAATTGAGATTT	CAGATTCAGATCGAGTTTG	CCCATTTTCGAGTTCAATGGTGCAC	719
Db	614	CCGTGG	AAGTTGAGATAT	TCAGAAATATGACGCTGTG	TGCATTTTGATCTTCAATAGCACCC	673
QY	720	CATTTC	CACATTCGATGATCTCT	CTCAATCCTTGAGGGGGTAA	AGCGGTAAACAGCATAGGAA	779
Db	674	CATTCT	GAGCTGACGACGACTCA	CTACGTCTTAAAGAAATGCGAT	TCTGTGGAGAGAAC	733
QY	780	GAGCTG	CGCCGC			792
Db	734	AACATG	CAGCAC			746

```

RESULT 7
US-10-425-115-2033
; Sequence 2033, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 2033
; LENGTH: 1483
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_101857C.1
US-10-425-115-2033

```



```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3170-045-C12_FLI
US-10-425-114-20051

Query Match      11.4%; Score 124.6; DB 17; Length 1167;
Best Local Similarity 57.4%; Pred. No. 1.7e-19;
Matches 245; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 332 CAAGGGTTTGATGAGAAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAAT 391
Db 273 CAACAATATGATGAGAAACATTCGTGGAAGGTGTATGATGCTTTGAATGTTCTCATG 332
QY 392 GCACCTTCGTTTATGCAAAAGAAAAAGAGAGATACGGTGGATGGCCCTTCCAAATTAC 451
Db 333 GCAATGGATATTATTTCTAAGGATAAAAAGGAAATTCATGGAAGGTCTCCCTCGTACT 392
QY 452 AGATATGAAAAATAAAGAGCTTTGAGGAATTCGTAAGAACTCGTCAACAGATTAGG 511
Db 393 AGTCTAAGCGATATTGAAGAGCTTAAAGTCGGAGCGTCTTGGGCTAAGGAATAGAA 452
QY 512 AACAAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAG 571
Db 453 AAGAAGCAGCCTATTGCAAGAGCTTAAAGTCGGAGCGTCTTGGGCTAAGGAATAGAA 512
QY 572 TTACGTAACCAACACTCGGAAAGCTTGAAGAGCTTGAAGAACTCGTCAACAGATTAGG 628
Db 513 CAACGAATGAGCAACTATATAGCTTAAAGTCGGAGCGTCTTGGGCTAAGGAATAGAA 572
QY 629 TTCGTATTGCTCAAGACATCTAGGAAGCAAGGTTGGAATTCAGATTTCCAGATGACTCG 688
Db 573 TTCAATCTGTGAGACACACGGCCCATGCAACTGTTGAAGTGAATATCAGAAATATG 632
QY 689 AAGTTGCCCATTCGAGTTCAATGTGTCACCAATTCACATTCGATGATGATCTCTCAATC 748
Db 633 CAGCTTGTCATTTGATTTCAATAGCACTCCCTTTGAGCTGATGATGATGATGATGTT 692
QY 749 CTTGAGG 755
Db 693 CTCAAGG 699

RESULT 10
US-10-424-599-43105
; Sequence 43105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43105
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_WRT3847_138923C.1
US-10-424-599-43105

Query Match      11.1%; Score 121.4; DB 17; Length 1751;
Best Local Similarity 56.9%; Pred. No. 1.2e-18;
Matches 243; Conservative 0; Mismatches 181; Indels 3; Gaps 1;

QY 332 CAAGGGTTTGATGAGAAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAAT 391
Db 715 CAACAATATGATGAGAAAAACATTCGTGGAAGGTGTATGATGCTCTGAAATGTTCTCATG 774
```

```
QY 392 GCACCTTCGTTGTTATTGCAAAAGAAAAAGAGATACGGTGGATGGGCTTTCAAATTAC 451
Db 775 GCATGGATATTATTCTTAAGGACAAAAGGAATTCATGGAAGGGTCTCCCTCGTACT 834
QY 452 AGATATGAAAAATAAAGAGCTTGAAGAGTTCGTAAGAACTCGTCAACAGATTAGG 511
Db 835 AGTCTAAGCGATATTGAAGAGCTTAAAGTCAGAGCGTCTTTGGGCTCAGGAATAGAA 894
QY 512 AACAAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAG 571
Db 895 AAGAAGCAGCCTATTTCGAAGAGCTGGAAGGCAATACATAGGTCTTCAGAAACTTATT 954
QY 572 TTACGTTAACCAACACTCGAAAGCTCAGCAGAGAATGTTAAT---GGCATCCGCTTCCA 628
Db 955 CAACGAATGAGCAACTATATAGCTCAGGAATGCTCCCAATGGAGGTGTATCTTGGCC 1014
QY 629 TTCGTATTGCTCAAGACATCTTAGGAAAGCAAGGTTGGAATTTGAGATTTAGATGACTCG 688
Db 1015 TTATTCTGTCAGACACGCGCCCATGCAACTGTTGAAGTGAATATCAGAAATATG 1074
QY 689 AAGTTGCCCATTCGAGTTCAATGTGTCACCAATTCACATTCGATGATGATCTCTCAATC 748
Db 1075 CAGCTTGTCATTTGATTATTTTGGCTCCCTTTGAGATGATGATGATGATGATGTT 1134
QY 749 CTTGAGG 755
Db 1135 CTCAAGG 1141

RESULT 11
US-10-425-114-36537
; Sequence 36537, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 36537
; LENGTH: 911
; TYPE: DNA
; ORGANISM: Zea mays subsp. mexicana
; FEATURE:
; OTHER INFORMATION: Clone ID: UC-ZMROTEOSINTE119B07_FLI
US-10-425-114-36537

Query Match      11.0%; Score 120; DB 17; Length 911;
Best Local Similarity 55.2%; Pred. No. 1.9e-18;
Matches 256; Conservative 0; Mismatches 205; Indels 3; Gaps 1;

QY 332 CAAGGGTTTGATGAGAAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTCTCAAT 391
Db 1 CAACAATATGATGAGAAAAATATACGACGAAGAGTTTATGATGCTTTGAATGTTCTGATG 60
QY 392 GCACCTTCGTTGTTATTGCAAAAGAAAAAGAGATACGGTGGATGGGCTTTCAAATTAC 451
Db 61 GCTATGGACATTATATCTAAAGATAAAAAGGAGATCCAGTGGAGGCTTTCGCGGTACT 120
QY 452 AGATATGAAAAATAAAGAGCTTGAAGAGTTGTAAGAGTTCGTAAGAACTCGTCAACAGATTAGG 511
Db 121 AGTATAAGTGACATTCGAAGAAATGAAGCGGAGCTTGTGGAGCTGAAGAGTAGGATTGAA 180
QY 512 AACAAAGAGGCACTCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAACATCAAG 571
Db 181 AAGAAAAGTCTTACCTCAGGAGCTACAAAGACCAATATGTAGTCTCGAAAAACCTGATT 240
```

QY 572 TTACGTAAACCAACACTGGAAGCTCAGCAGA---GAATGTTAATGTCATCGGCTTCCA 628
Db 241 CAACGAAATGAGCAATCATATGTTTCAGGAAACACGCTTCTGCTGAGTGGCTTTGCCA 300
QY 629 TTCGTATTGTCAGACATCTAGGAAGCAGGCTGGAATTTAGATTTAGATGACTCG 688
Db 301 TTTATCCTAGTTTCAGACCCGACCTCATGCTACTGTGGAAGTTGAAATATCAGAAGATATG 360
QY 689 AAGTTTCCCATTTTCGAGTTCAATGCTGCACCTTACATTTGATGATGATCTCTCAATC 748
Db 361 CAGCTGTGATTTTGAATTTCAATAGACCCCTTCAGCTGCAATGACGACTCATAGCTC 420
QY 749 CTTGAGGGGTAAAGCGGTAAACAGCATAGGAGAGCTGCGCGCG 792
Db 421 CTAAGAAGAAATGCGATCTTCTGTGGAAGAGAACACATGACAGCAC 464

RESULT 12
US-10-767-795-6106
; Sequence 6106, Application US/10767795
; Publication No. US2004018130A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Cao, Yongwei
; APPLICANT: Zhou, Yihua
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53534)B
; CURRENT APPLICATION NUMBER: US/10/767,795
; CURRENT FILING DATE: 2004-01-30
; NUMBER OF SEQ ID NOS: 117596
; SEQ ID NO 6106
; LENGTH: 642
; TYPE: DNA
; ORGANISM: Gossypium hirsutum
; FEATURE:
; OTHER INFORMATION: Clone ID: GOSHI-09MAY01-C6102_1
US-10-767-795-6106

Query Match 9.4%; Score 102.8; DB 18; Length 642;
Best Local Similarity 54.9%; Pred. No. 2.1e-14;
Matches 225; Conservative 0; Mismatches 182; Indels 3; Gaps 1;
QY 332 CAAGGGTTTTCATGAGAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTTTCATT 391
Db 85 CAACAGTTCAGCAGAGAAACATCCGCCGAGGGTTTATGATGCTCTGAATGTTCTAATG 144
QY 392 GCACTTCGTGTTATTGCAAGAAAAAGGAGATACGGTGGATGGCCCTTTCAAAATTAC 451
Db 145 GCAATGGATATTATATCTAAGGATAAAAAAGAAATACAGGGAGGGACGGCTTCGTACT 204
QY 452 AGATATGAAAAATAAGAAAGCTTGAGGAATGTCGTAAGAACTCGTCAACAGATAGG 511
Db 205 GGCTCAGTGATATTGAAGATTTTAAAGACTGAACGGCTTGGGCTGAGAAATAGGATTGAT 264
QY 512 AACAGAAAGCACTCTCCAGGAAATCGAAACACAGTTTGTATGATCTCCAAACATCAAG 571
Db 265 AAAAAAGCTGCCTATTGTGATGTAATTTGAGGAACAGTTTGTAGTCTTCAGAAATTTGATA 324
QY 572 TTACGTAAACCAACACTGGAAGCTCAGCAGAGAAT---GTTAATGGCATCCGCTTCCA 628
Db 325 CAACGAAATGAGCAACTATACAGCTCAGGCAATGCTTCCGGTGGGGAGTGGCTTTACCT 384
QY 629 TTCGTATTGTCAGACATCTAGGAAGCAGGCTGGAATTTAGATTTTCAAGATGACTCG 688
Db 385 TTTATCCTGTCAGACACGCGCTCATGCAACTGTTGAGTAGAAATATCGGAAGATATG 444
QY 689 AAGTTTCCCATTTTCGAGTTCAATGCTGCACCTTACATTTGATGATCA 738
Db 445 CAGCTGTGCAATTTGACTTTAATAGACTTCCCTTTGAGCTCCAGATGA 494

RESULT 13
US-10-198-846-10136
; Sequence 10136, Application US/10198846
; Publication No. US2003009974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steimann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10136
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1_2
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10136

Query Match 9.1%; Score 99.2; DB 14; Length 2968;
Best Local Similarity 58.2%; Pred. No. 3.2e-13;
Matches 199; Conservative 0; Mismatches 133; Indels 10; Gaps 1;
QY 236 ATAGTTTGTGAGAAAGTTGAAGCCAAAGGAAGAACATACATAGGTTGAGAGCAA 295
Db 403 AAAGTGTGTGAGAAAGTTTCAACGAAAGGTACACATCGTACATGAAGTTCGTATGAG 462
QY 296 ATTTATTTCAGAGCTGAAGT-----CCATGGCACATATTGGTCAAGGGTTTGATGA 345
Db 463 CTGGTGTCTAGAGTTTCAACATTAACCAATTTGGTGTCTGCTGATTCGGCTTATGATCA 522
QY 346 GAAGATATTAGCGGAGAGTGTATGATGCTTTCAACGTTTCTCATTCGCTGCTAT 405
Db 523 GAAGAACATTTAGCGGAGAGTTTATGATGCTTTAATGTCTAATGCAATGAACATAAT 582
QY 406 TGCAGAAAGAAAAAGGAGATACGGTGGATGGCCCTTTCAAATTCAGATATGAAAAAT 465
Db 583 TTCAAGGAAAAAGAAATCAAGTGGATTTGGCTTCCCTTACCAATTTCTGCTCAGGAATG 642
QY 466 AAAGAGCTTTGAGGAGTTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACT 525
Db 643 TCAGATCTCGAGATAGAGAACAGAGGCGGATAGAACGGATTAACAGAACGGCGCCA 702
QY 526 CTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAAACAT 567
Db 703 GCTGCAAGAACTTCTCTACAGCAAAATCGTTTCAAAAACCT 744

RESULT 14
US-10-641-643-1456
; Sequence 1456, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; TITLE OF INVENTION: GENE EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO

STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641.643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1456:
SEQUENCE CHARACTERISTICS:
LENGTH: 2320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: G604478
SEQUENCE DESCRIPTION: SEQ ID NO: 1456 :
US-10-641-643-1456

Query Match 9.1%; Score 98.8; DB 17; Length 2320;
Best Local Similarity 58.1%; Pred. No. 3.6e-13;
Matches 200; Conservative 0; Mismatches 132; Indels 12; Gaps 1;
QY 236 ATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGTTGCGAGCGAA 295
DB 369 AAAGTGTGTGAGAAAGTTCAACGAAAGGTACACATCGTACAAATGAAAGTCGCTGATGAG 428
QY 296 ATTATTTCAGAGCTG-----AAGTCCATGGCACATATTGGTCAAGGGTTTGAT 343
DB 429 CTGGTGTGAGAGTTTCAACAAATTCAAATACCAATTTGGCTGCTGATTCGAGGCTTAATGAT 488
QY 344 GAGAAGAAATATTAGCGGAGAGTGTATGATGCTTTCAACGCTTCTCAATTGCACTTCGTTGTT 403
DB 489 CAGAAGAACATTAGCGGAGAGTTTATGATGCTTTAAATGTGCTAATGGCAATGAACATA 548
QY 404 ATTGCAAAAGAAAAGGAGATACGGTGGATGGCCCTTTCAAAATACAGATATGAAAAA 463
DB 549 ATTTCAAAGGAAAAGAAAATCAAGTGGATTGGCCCTGCTACCAATTTCTGCTCAGGAA 608
QY 464 ATRAAGAGCTTGAGGAAGTTTCGTAAGAACTCGTCAACAGATATTAGGAACAAGAGGCA 523
DB 609 TGTCAAGATCTGGAGATAGAGAACAGAGCGGATAGACGGATAGAGAGAGCGGCC 668
QY 524 CTCCTCCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAACAT 567
DB 669 CAGCTGCAAGAACTTCTCCTACAGCAAAATCGCTTTCAAAAACCT 712

RESULT 15
US-09-954-456-1176
; Sequence 1176, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Cand
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76
; CURRENT APPLICATION NUMBER: US/09/954.456

CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233.617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234.052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234.923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235.134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235.637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235.638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235.711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235.720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235.840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235.863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1176
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1176

Query Match 8.9%; Score 97; DB 9; Length 1266;
Best Local Similarity 57.8%; Pred. No. 7.3e-13;
Matches 197; Conservative 0; Mismatches 135; Indels 9; Gaps 1;
QY 236 ATAGTTGTGAGAAAGTTGAAGCCAAAGGAAGAACACATACATGAGTTGCGAGCGAA 295
DB 261 AAAGTGTGTGAGAAAGTTCAACGAAAGGTACACATCGTACAAATGAAAGTCGCTGATGAG 320
QY 296 ATTATTTCAGAGCTG-----AAGTCCATGGCACATATTGGTCAAGGGTTTGATGAG 346
DB 321 CTGGTGTGAGAGTTTCAACAAATTCAAATACCAATTTGGCTGCTGATTCGCTTATGATCAG 380
QY 347 AAGAATATTAGCGGAGAGTGTATGATGCTTTCAACGTTTCTCATGCACTTCGTTGTTATT 406
DB 381 AAGAATATTAGCGGAGAGTGTATGATGCTTTAAATGTCTAATGCAATGAACATAATT 440
QY 407 GCAAGAAAAAGAGAGATACGGTGGATGGCCCTTTCAAAATACAGATATGAATAATA 466
DB 441 TCAAGAGAAAAAGAAATCAAGTGGATTGGCTGCTACCAATTTCTGCTCAGGAATGT 500
QY 467 AAGAGCTTTGAGGAAGTTTCGTAAGAACTCGTCAACAGATTAGGAACAAGAGGCACTC 526
DB 501 CAGATCTGAGATAGAGAGCAGAGCGCGATAGAACGGATAGAACGAGAGCGGCCAG 560
QY 527 CTCAGGAAATCGAAAAACAGTTTGTATGATCTCCAAAACAT 567
DB 561 CTGCAAGAACTTCTCCTACAGCAAAATCGCTTTCAAAAACCT 601

Search completed: February 6, 2005, 17:24:00
Job time : 693 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2005, 20:05:19 ; Search time 165 Seconds
(without alignments)
611.785 Million cell updates/sec

Title: US-10-088-830-2
Perfect score: 1319
Sequence: 1 MAPPRGGAATAALDLTG.....SILEGVRNSIGRAGRATLH 261

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1314	99.6	261	4	AAB67762
2	890.5	67.5	294	8	ADQ36843
3	492	37.3	344	3	AAG22851
4	492	37.3	385	3	AAG22850
5	492	37.3	385	5	AU72542
6	492	37.3	385	8	ADO62143
7	492	37.3	385	8	ADO63083
8	471.5	35.7	346	8	ADQ36845
9	466.5	35.4	376	5	AU72514
10	466.5	35.4	379	8	ADQ36855
11	461	35.0	299	3	AAB07975
12	437	33.1	277	3	AAG22852
13	428.5	32.5	292	5	AU72529
14	428.5	32.5	292	6	ABP96848
15	428.5	32.5	292	7	ADH59560
16	428.5	32.5	292	8	ADO63085
17	428.5	32.5	292	8	ADO62145
18	428.5	32.5	296	5	AU72497
19	416.5	31.6	194	2	AAY32164
20	412	31.2	262	5	AU72572
21	406.5	30.8	251	5	AU72559
22	399.5	30.3	431	8	ADN05730
23	396.5	30.1	424	4	ABG28057
24	389	29.5	446	2	AAY18026
25	378.5	28.7	410	2	AAR89212

ALIGNMENTS

RESULT 1

AAB67762
ID AAB67762 standard; protein; 261 AA.

XX AC AAB67762;

DT XX AC

DT 11-JUN-2001 (first entry)

DE Amino acid sequence of a wheat E2F-dimerisation partner (DP) protein.

XX E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;

KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation.

XX OS Triticum monococcum.

XX FN WO200121644-A2.

XX PD 29-MAR-2001.

XX PF 25-SEP-2000; 2000WO-EP009325.

XX PR 24-SEP-1999; 99ES-00002127.

XX PR 11-NOV-1999; 99ES-00002474.

XX (CNSJ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.

XX Gutierrez-Armenta C, Ramirez-Parra E;

XX WPI; 2001-257972/26.

XX N-PSDB; AAF80144.

PT New isolated, enriched, cell free and/or recombinant nucleic acid useful for e.g. altering cell proliferation characteristic such as to alter plant cell, organ or tissue size.

PT Claim 7; Fig 1; 77pp; English.

PS The present sequence represents a E2F-dimerisation partner (DP) protein. The protein acts as a plant E2F transcription factor. E2F and DP are two proteins that hetero-dimerise to form an active transcription factor that regulates G1 to S phase of the cell cycle, and later, the expression of genes required for S-phase progression. E2F and retinoblastoma protein also interact as a hetero-dimer in cells to suppress certain genes. This repression involves binding of the retinoblastoma protein to the E2F-DP dimer that is in turn bound to sites on DNA through the E2F DNA binding domain. DP proteins can be modulated to alter plant cell, organ or tissue shape, and particularly to alter cell proliferation characteristic such as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

CC as to alter plant cell, organ or tissue size

XX SQ Sequence 261 AA;
Query Match 99.6%; Score 1314; DB 4; Length 261;
Best Local Similarity 99.6%; Pred. No. 5.6e-115;
Matches 260; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MAPPRGAAAAATAALDLTGWHILEASSVPPPLPEAGNAVORVGAVDPDKDRKKEKAAAP 60
DB 1 MAPPRGAAAAATAALDLTGWHILEASSVPPPLPERGNAVORVGAVDPDKDRKKEKAAAP 60
QY 61 RITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMAHIGQGFEKNTIRRRVYDAF 120
DB 61 RITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMAHIGQGFEKNTIRRRVYDAF 120
QY 121 NVLIARLVIKKEIRWMLGSLNRYEYKIKLEVRKELVKNIRKALLQETIEKQFDDL 180
DB 121 NVLIARLVIKKEIRWMLGSLNRYEYKIKLEVRKELVKNIRKALLQETIEKQFDDL 180
QY 181 ONIKLRNQTLESSAENVNGIRLPVLVKTSRKARVEIEISDDSKFAHFFENGAPFTLHDD 240
DB 181 ONIKLRNQTLESSAENVNGIRLPVLVKTSRKARVEIEISDDSKFAHFFENGAPFTLHDD 240
QY 241 LSILEGVRNSIGRAGRATLH 261
DB 241 LSILEGVRNSIGRAGRATLH 261
RESULT 2
ADQ36843
ID ADQ36843 standard; protein; 294 AA.
AC ADQ36843;
XX
DT 07-OCT-2004 (first entry)
DE
DE Os018989-4003 protein.
XX cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response.
XX
OS Oryza sativa.
XX
XX W02004061122-A2.
XX
PD 22-JUL-2004.
XX
XX 23-DEC-2003; 2003WO-US041200.
XX
XX 26-DEC-2002; 2002US-0436565P.
XX
PA (SYGN) SYNGENTA PARTICIPATIONS AG.
XX
XX Cooper B;
XX
XX WPI; 2004-534388/51.
XX
XX New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX
XX Claim 28; SEQ ID NO 2; 409pp; English.
XX
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related polypeptide. The
CC present sequence is published separately from the main body of the
CC specification as EPO data.
XX

SQ Sequence 294 AA;
Query Match 67.5%; Score 890.5; DB 8; Length 294;
Best Local Similarity 63.1%; Pred. No. 4.7e-75;
Matches 185; Conservative 31; Mismatches 36; Indels 41; Gaps 4;
QY 1 MAPPRGAAAAATAALDLTGWHILEASSVPPPLPE----- 34
DB 1 MAPPCGDAAAAASAAPGLNLLIRGAGLPSPERYPPRPCTSDSFAPISREGDDIPPQ 60
QY 35 -----AGNAVORK--GAVDPDKDRKKEKAAAPRITGWLREYSKIVCEKVEAKGR 83
DB 61 KKSVSILRSGGGGNAAREEGGA---NRNGKKEKTAQRTITGWLREFSKIVSKVEAKGR 117
QY 84 TTYNEVADEIYSELKSMAHIGQGFEKNTIRRRVYDAFNVLIARLVIKKEIRWMLGSLN 143
DB 118 TTYNEVADEIYFAELKSIITQNGLEFDEKNTIRRRVYDAFNVLIARLVIKKEIRWMLGSLN 177
QY 144 YYEYKIKLEVRKELVKNIRKALLQETIEKQFDDLQNIKLNRNQTLESSAENVNGIRLP 203
DB 178 YYEYKIKLEVRKELVKNIRKALLQETIEKQFDDLQNIKLNRNQTLESSAENVNGIRLP 237
QY 204 FVLVKTSRKARVEIEISDDSKFAHFFENGAPFTLHDDLSILEGVRNSIGRAG 256
DB 238 FLIIKTSRKARVEIEISDDSKFAHFFENGAPFTLHDDLSILEGVRNSIGRAG 289
RESULT 3
AAG22851
ID AAG22851 standard; protein; 344 AA.
AC AAG22851;
XX
DT 17-OCT-2000 (first entry)
DE
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25936.
XX Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 29-MAR-1999; 99US-0126264P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.
 PR 14-MAY-1999; 99US-0134221P.
 PR 14-MAY-1999; 99US-0134370P.
 PR 18-MAY-1999; 99US-0134768P.
 PR 19-MAY-1999; 99US-0134941P.
 PR 20-MAY-1999; 99US-0135124P.
 PR 21-MAY-1999; 99US-0135353P.
 PR 24-MAY-1999; 99US-0135629P.
 PR 25-MAY-1999; 99US-0136021P.
 PR 27-MAY-1999; 99US-0136392P.
 PR 28-MAY-1999; 99US-0136782P.
 PR 01-JUN-1999; 99US-0137222P.
 PR 03-JUN-1999; 99US-0137528P.
 PR 04-JUN-1999; 99US-0137502P.
 PR 07-JUN-1999; 99US-0137724P.
 PR 08-JUN-1999; 99US-0138094P.
 PR 10-JUN-1999; 99US-0138540P.
 PR 10-JUN-1999; 99US-0138847P.
 PR 14-JUN-1999; 99US-0139119P.
 PR 16-JUN-1999; 99US-0139452P.
 PR 17-JUN-1999; 99US-0139453P.
 PR 17-JUN-1999; 99US-0139492P.
 PR 18-JUN-1999; 99US-0139454P.
 PR 18-JUN-1999; 99US-0139455P.
 PR 18-JUN-1999; 99US-0139456P.
 PR 18-JUN-1999; 99US-0139457P.
 PR 18-JUN-1999; 99US-0139458P.
 PR 18-JUN-1999; 99US-0139459P.
 PR 18-JUN-1999; 99US-0139460P.
 PR 18-JUN-1999; 99US-0139461P.
 PR 18-JUN-1999; 99US-0139462P.
 PR 18-JUN-1999; 99US-0139463P.
 PR 18-JUN-1999; 99US-0139750P.
 PR 18-JUN-1999; 99US-0139763P.
 PR 21-JUN-1999; 99US-0139817P.
 PR 22-JUN-1999; 99US-0139899P.
 PR 23-JUN-1999; 99US-0140353P.
 PR 23-JUN-1999; 99US-0140354P.
 PR 24-JUN-1999; 99US-0140695P.
 PR 28-JUN-1999; 99US-0140823P.
 PR 29-JUN-1999; 99US-0140991P.
 PR 30-JUN-1999; 99US-0141287P.
 PR 01-JUL-1999; 99US-0141842P.
 PR 01-JUL-1999; 99US-0142154P.
 PR 02-JUL-1999; 99US-0142055P.
 PR 06-JUL-1999; 99US-0142390P.
 PR 08-JUL-1999; 99US-0142803P.
 PR 09-JUL-1999; 99US-014220P.
 PR 12-JUL-1999; 99US-014277P.
 PR 13-JUL-1999; 99US-0143542P.
 PR 14-JUL-1999; 99US-0143624P.
 PR 15-JUL-1999; 99US-0144005P.
 PR 16-JUL-1999; 99US-0144085P.
 PR 16-JUL-1999; 99US-0144086P.
 PR 19-JUL-1999; 99US-0144325P.
 PR 19-JUL-1999; 99US-0144331P.
 PR 19-JUL-1999; 99US-0144332P.
 PR 19-JUL-1999; 99US-0144333P.
 PR 19-JUL-1999; 99US-0144334P.
 PR 19-JUL-1999; 99US-0144335P.
 PR 20-JUL-1999; 99US-0144352P.
 PR 20-JUL-1999; 99US-0144632P.
 PR 20-JUL-1999; 99US-0144884P.
 PR 21-JUL-1999; 99US-0144814P.
 PR 21-JUL-1999; 99US-0145086P.
 PR 21-JUL-1999; 99US-0145088P.
 PR 22-JUL-1999; 99US-0145085P.
 PR 22-JUL-1999; 99US-0145087P.
 PR 22-JUL-1999; 99US-0145089P.
 PR 22-JUL-1999; 99US-0145192P.
 PR 23-JUL-1999; 99US-0145145P.
 PR 23-JUL-1999; 99US-0145218P.
 PR 23-JUL-1999; 99US-0145224P.

PR 26-JUL-1999; 99US-0145276P.
 PR 27-JUL-1999; 99US-0145913P.
 PR 27-JUL-1999; 99US-0145918P.
 PR 27-JUL-1999; 99US-0145919P.
 PR 28-JUL-1999; 99US-0145951P.
 PR 02-AUG-1999; 99US-0146386P.
 PR 02-AUG-1999; 99US-0146388P.
 PR 02-AUG-1999; 99US-0146389P.
 PR 03-AUG-1999; 99US-0147038P.
 PR 04-AUG-1999; 99US-0147204P.
 PR 04-AUG-1999; 99US-0147302P.
 PR 05-AUG-1999; 99US-0147192P.
 PR 05-AUG-1999; 99US-0147260P.
 PR 06-AUG-1999; 99US-0147303P.
 PR 06-AUG-1999; 99US-0147416P.
 PR 09-AUG-1999; 99US-0147493P.
 PR 09-AUG-1999; 99US-0147935P.
 PR 10-AUG-1999; 99US-0148171P.
 PR 11-AUG-1999; 99US-0148319P.
 PR 12-AUG-1999; 99US-0148341P.
 PR 13-AUG-1999; 99US-0148565P.
 PR 13-AUG-1999; 99US-0148684P.
 PR 16-AUG-1999; 99US-0149368P.
 PR 17-AUG-1999; 99US-0149175P.
 PR 18-AUG-1999; 99US-0149426P.
 PR 20-AUG-1999; 99US-0149722P.
 PR 20-AUG-1999; 99US-0149723P.
 PR 20-AUG-1999; 99US-0149929P.
 PR 23-AUG-1999; 99US-0149902P.
 PR 23-AUG-1999; 99US-0149930P.
 PR 25-AUG-1999; 99US-0150566P.
 PR 26-AUG-1999; 99US-0150884P.
 PR 27-AUG-1999; 99US-0151065P.
 PR 27-AUG-1999; 99US-0151066P.
 PR 27-AUG-1999; 99US-0151080P.
 PR 30-AUG-1999; 99US-0151303P.
 PR 31-AUG-1999; 99US-0151438P.
 PR 01-SEP-1999; 99US-0151930P.
 PR 07-SEP-1999; 99US-0152363P.
 PR 10-SEP-1999; 99US-0153070P.
 PR 13-SEP-1999; 99US-0153758P.
 PR 15-SEP-1999; 99US-0154018P.
 PR 16-SEP-1999; 99US-0154039P.
 PR 20-SEP-1999; 99US-0154779P.
 PR 22-SEP-1999; 99US-0155139P.
 PR 23-SEP-1999; 99US-0155486P.
 PR 24-SEP-1999; 99US-0155659P.
 PR 28-SEP-1999; 99US-0156458P.
 PR 29-SEP-1999; 99US-0156596P.
 PR 04-OCT-1999; 99US-0157117P.
 PR 05-OCT-1999; 99US-0157753P.
 PR 06-OCT-1999; 99US-0157865P.
 PR 07-OCT-1999; 99US-0158029P.
 PR 08-OCT-1999; 99US-0158232P.
 PR 12-OCT-1999; 99US-0158369P.
 PR 13-OCT-1999; 99US-0159293P.
 PR 13-OCT-1999; 99US-0159294P.
 PR 13-OCT-1999; 99US-0159295P.
 PR 14-OCT-1999; 99US-0159329P.
 PR 14-OCT-1999; 99US-0159330P.
 PR 14-OCT-1999; 99US-0159331P.
 PR 14-OCT-1999; 99US-0159637P.
 PR 14-OCT-1999; 99US-0159638P.
 PR 18-OCT-1999; 99US-0159584P.
 PR 21-OCT-1999; 99US-0160741P.
 PR 21-OCT-1999; 99US-0160767P.
 PR 21-OCT-1999; 99US-0160768P.
 PR 21-OCT-1999; 99US-0160770P.
 PR 21-OCT-1999; 99US-0160814P.
 PR 21-OCT-1999; 99US-0160815P.
 PR 22-OCT-1999; 99US-0160980P.
 PR 22-OCT-1999; 99US-0160981P.
 PR 22-OCT-1999; 99US-0160989P.


```
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145291P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147392P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160981P.
PR 25-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 37.3%; Score 492; DB 3; Length 385;
Best Local Similarity 47.7%; Pred. No. 2.2e-37;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVQKGAVDPKDKKKKAAAPRITGWLREYSKIVCEKAEKGRITTYNEVADITYSE 96
Db 74 GDDAGSQGASGVKKKRGQRAAGPKTGRGLKQFSMKVCEKESKRTTYNEVADLVAE 133
QY 97 LKSMAHIG----QGDEKNIRRRVYDAFNVLIALRVIAKEKKEIRRMGLSNRYRKIKK 151
Db 134 FALPNNDGTSPOQOQYDEKNIRRRVYDALNVLMDMDIISKDKKEIQWRGLPRTSLSDIEE 193
QY 152 LEEVRKELVKNIRKKALLQETKEQDPDDLQNLKRNQTLLESSAENVN-GIRLPFVLVKT 210
Db 194 LKNERLSLRNRIEKTAYSQELEEQYVGLQNLQORNEHLYSSGNAPSGGVLPFILVQTR 253
QY 211 RKARVEIEISDDSKFAHFEFNGAPFTLHDDLSILEGVR 248
Db 254 PHATVEVEISEDMQLVHFDNFSTPFELHDDNFVLKTMK 291

RESULT 5
AAU72542
ID AAU72542 standard; protein; 385 AA.
AC AAU72542;
XX
XX
DT 26-FEB-2002 (first entry)
XX
DE Arabidopsis cell cycle protein CCP33.
KW Cell cycle protein; CCP; cell cycle regulation; herbicide;
KW plant growth regulator; plant development; abiotic stress; biotic stress;
KW nutrient deprivation; pathogen attack; crop yield; immunogen.
XX
OS Arabidopsis thaliana.
XX
XX WO200185946-A2.
XX
PD 15-NOV-2001.
XX
PF 14-MAY-2001; 2001WO-IB001307.
XX
PR 12-MAY-2000; 2000US-0204045P.
XX
PA (CROP-) CROPDESIGN NV.
PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;
XX
DR WPI; 2002-062249/08.
DR N-PSDB; AAS96332.
XX
PT New cell cycle protein and nucleic acid molecule encoding it useful for
PT regulating cell cycle progression in plants and for identifying
PT modulators which are useful as herbicides or plant growth regulators.
XX
XX Claim 34; Fig 39; 316pp; English.
```

```

XX CC The invention relates to a novel cell cycle protein (CCP) and the
XX CC polynucleotides encoding them. CCP is useful for identifying a compound
XX CC which modulates the activity of the polypeptide and which binds to the
XX CC polypeptide and an anti-CCP antibody is useful for detecting the presence
XX CC of CCP in a sample. A CCP modulator is useful for modulating the cell
XX CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
XX CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP
XX CC nucleic acid and polypeptide molecules are useful as modulating agents in
XX CC regulating cell cycle progression in plants. CCP is useful to treat
XX CC disorders characterised by insufficient or excessive production of CCP
XX CC protein or production of CCP protein forms which have decreased or
XX CC aberrant activity. Compounds that bind to or modulate the activity of CCP
XX CC polypeptide are useful as herbicides or plant growth regulators. The
XX CC polynucleotide is useful for modifying cell fate, plant development,
XX CC plant morphology, biochemistry and/or physiology, the length of the G1,
XX CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
XX CC stimulation or enhancement of cell division, DNA replication, seed set,
XX CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
XX CC initiation and/or development, nodule function, dwarfism in plants,
XX CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
XX CC and the anti-CCP antibody are useful in agriculture to modulate the
XX CC protein levels or activity of a protein involved in the cell cycle due to
XX CC environmental conditions, including abiotic stresses such as cold, nutrient
XX CC deprivation, heat, drought, salt stress, or biotic stress such as
XX CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate
XX CC plant architecture, plant quality traits, plant reproduction and seed
XX CC development, endoreduplication in storage cells, storage tissues and/or
XX CC storage organs of plants or its parts. CCP is useful as an immunogen to
XX CC generate antibodies. CCP protein is useful to screen for naturally
XX CC occurring CCP substrates. The polynucleotide is useful for expressing CCP
XX CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to
XX CC modulate CCP activity. The present sequence represents a CCP protein of
XX CC the invention
XX SQ Sequence 385 AA;

Query Match 37.3%; Score 492; DB 5; Length 385;
Best Local Similarity 47.7%; Pred. No. 2.2e-37;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVQRKGA VDPDKRKKAAPRITGMGLREYSKIVCEKBAKGRTTYNEVADIYSE 96
DB 74 GDDAGSQGASGVKKRGQRAAGPDKTGRGLRQFSMKVCEKVBKGRTTYNEVADLVAE 133
QY 97 LKSMAHIG-----QGFEKNIRRRVYDAFNVLIALRVIAKEKKEIRWMLSNRYEKIKK 151
DB 134 FALPNNDDGTSPDQQYDEKNIRRRVYDALNVLNMDIISKDKKEIQWRGLPRTSLSDIER 193
QY 152 LEEVRKELVNKIRNKALQIEKQFDLQNIKLKRNQTLLESSAENVN-GIRLPFVLKTS 210
DB 194 LKNERLSLRNIRBKKTAYSQLEBEQYVGLQNLQORNEHLIYSSGNAPSGGVALPFIIVQTR 253
QY 211 RKARVIEISDDSKFAHFEPNGAPFTLHDLISLEGVR 248
DB 254 PHATVEISEDMQLVHFDNFTSPFELHDDNFVLKTMK 291

RESULT 6
AD062143
ID AD062143 standard; protein; 385 AA.
XX AC ADO62143;
XX DT 15-JUL-2004. (first entry)
XX DE Transcription factor G2981, SEQ ID 610.
XX KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
XX KW osmotic stress tolerance; cold tolerance; heat tolerance;
XX KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
XX KW glyphosate resistance; flowering; fertility; seed development.

```

```

OS Arabidopsis thaliana.
XX PN WO20040311349-A2.
XX PD 15-APR-2004.
XX PF 18-SEP-2003; 2003WO-US030292.
XX PR 18-SEP-2002; 2002US-0411837P.
XX PR 17-DEC-2002; 2002US-0434166P.
XX PR 24-APR-2003; 2003US-0465809P.
XX PA (WEND-) MENDEL BIOTECHNOLOGY INC.
XX PI Jiang C, Heard JF, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL,
PI Riechmann JL, Haake V, Dubell JS, Sherman BK;
XX DR WPI; 2004-330163/30.
XX DR N-PSDB; ADO62142.
XX PT New recombinant polynucleotide encoding transcription factor
XX PT polypeptides, useful for producing transgenic plants with advantageous
XX PT properties compared to a reference plant.
XX PS Claim 14; SEQ ID NO 610; 510pp; English.
XX CC The present invention relates to novel plant transcription factor
XX CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
XX CC sequences can be used to produce transgenic plants, which overexpress
XX CC (II), where the transgenic plant has an altered trait as compared to a
XX CC non-transgenic plant or wild-type plant. The transgenic plant comprises
XX CC an altered trait selected from increased tolerance to abiotic stress,
XX CC increased tolerance to osmotic stress, increased tolerance to cold,
XX CC increased germination in cold, increased tolerance to heat, increased
XX CC germination in heat, increased tolerance to freezing conditions,
XX CC increased tolerance to low nitrogen conditions, increased tolerance to
XX CC low phosphate conditions, increased tolerance to disease, including
XX CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
XX CC increased tolerance to multiple fungal pathogens, increased resistance to
XX CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
XX CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
XX CC to sugars, altered carbon/nitrogen sensing, early flowering, late
XX CC flowering, altered flower structure, loss of flower determinacy, reduced
XX CC fertility, altered shoot meristem development, altered branching pattern,
XX CC altered stem morphology, altered vascular tissue structure, reduced
XX CC apical dominance, altered trichome density, altered trichome development,
XX CC altered trichome structure, altered root development, altered shade
XX CC avoidance, altered seed development, altered seed ripening, altered seed
XX CC germination, slow growth, fast growth, altered cell differentiation,
XX CC altered cell proliferation, altered cell expansion, altered programmed cell
XX CC death, lethality when overexpressed, altered necrosis patterns, increased
XX CC plant size, increased biomass, large seedlings, dwarfed plants, dark
XX CC green leaves, change in leaf shape, increased leaf size and mass, light
XX CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
XX CC altered seed coloration, altered seed size, altered seed shape, large
XX CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
XX CC content, altered seed protein content, altered seedprenyl content,
XX CC altered leaf prenyl lipid content, increased anthocyanin levels, and
XX CC decreased anthocyanin levels. Note: The sequence data for this patent did
XX CC not form part of the printed specification, but was obtained in
XX CC electronic format directly from WIPO at
XX CC ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 385 AA;

Query Match 37.3%; Score 492; DB 8; Length 385;
Best Local Similarity 47.7%; Pred. No. 2.2e-37;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;

QY 37 GNAVQRKGA VDPDKRKKAAPRITGMGLREYSKIVCEKBAKGRTTYNEVADIYSE 96
DB 74 GDDAGSQGASGVKKRGQRAAGPDKTGRGLRQFSMKVCEKVBKGRTTYNEVADLVAE 133

```

QY 97 LKSMHIG-----QGDEKNIRRVYDAFNVLIALRVIAKEKEIRWMGLSNRYEKKK 151
Db 134 FALPNDGTSPDQQYDEKNIRRVYDALNVLNMDIISKKEIQWRGLPRTSLSDIEE 193
QY 152 LEEVRKELVNKRNKALLQIEIKQFDDIQLNIKLNRQTLESSAENVN-GIRLPFVLVKT 210
Db 194 LKNERLSLRNRIEKTAYSQELEEQVGLQNLQORNEHLYSSGNAPSGGVALPFILVQTR 253
QY 211 RKARVEIETSDSKFAHFEFNGAPFTLHDDLSTLEGVR 248
Db 254 PHATVEVEISEDMQLVHDFNSTPFELHDDNFVLKTMK 291

RESULT 7
AD063083
ID ADO63083 standard; protein; 385 AA.
XX ADO63083;
DT 15-JUL-2004 (first entry)
DE Transcription factor G2981, SEQ ID 1550.
XX Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KW osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development.
XX Arabidopsis thaliana.
OS
XX WO2004031349-A2.
PN
XX 15-APR-2004.
PD
XX 18-SEP-2003; 2003WO-US030292.
PF
XX 18-SEP-2002; 2002US-0411837P.
PR
XX 17-DEC-2002; 2002US-0434166P.
PR
XX 24-APR-2003; 2003US-0465809P.
PR
XX (MEND-) MENDEL BIOTECHNOLOGY INC.
PA
XX Jiang C, Heard JE, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JB, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX
XX WPI; 2004-330163/30.
DR
XX N-PSDB; ADO63082.
DR
XX New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
XX Claim 1; SEQ ID NO 1550; 510pp; English.
PS
XX The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-A063778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,

CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 385 AA;

Query Match 37.3%; Score 492; DB 8; Length 385;
Best Local Similarity 47.7%; Pred. No. 2.2e-37;
Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2;
QY 37 GNAVQRKGAVDPDKRKKEKAAAPRITGWLREYSKIVCEKVENKGRTTYNEVADIYSE 96
Db 74 GDDAGSQGASGVKKKRGQRAAGPKTGRGLRQFSMKVCEKVESKGRTTYNEVADSLVAE 133
QY 97 LKSMHIG-----QGDEKNIRRVYDAFNVLIALRVIAKEKEIRWMGLSNRYEKKK 151
Db 134 FALPNDGTSPDQQYDEKNIRRVYDALNVLNMDIISKKEIQWRGLPRTSLSDIEE 193
QY 152 LEEVRKELVNKRNKALLQIEIKQFDDIQLNIKLNRQTLESSAENVN-GIRLPFVLVKT 210
Db 194 LKNERLSLRNRIEKTAYSQELEEQVGLQNLQORNEHLYSSGNAPSGGVALPFILVQTR 253
QY 211 RKARVEIETSDSKFAHFEFNGAPFTLHDDLSTLEGVR 248
Db 254 PHATVEVEISEDMQLVHDFNSTPFELHDDNFVLKTMK 291

RESULT 8
AD036845
ID ADQ36845 standard; protein; 346 AA.
XX ADQ36845;
AC ADQ36845;
DT 07-OCT-2004 (first entry)
XX Cell proliferation-related polypeptide #1.
DE Cell proliferation-related polypeptide; cell proliferation; senescence;
XX cell proliferation related polypeptide; cell proliferation; stress response;
KW differentiation; stress response.
OS Oryza sativa.
XX WO2004061122-A2.
PN
XX 22-JUL-2004.
PD
XX 23-DEC-2003; 2003WO-US041200.
PF
XX 26-DEC-2002; 2002US-0436565P.
PR
XX (SYGN) SYNGENTA PARTICIPATIONS AG.
PA
XX Cooper B;
PI
XX WPI; 2004-534388/51.
DR
XX New nucleic acid molecule encoding a cell proliferation-related
PT

KW cell proliferation related polypeptide; cell proliferation; senescence;
XX differentiation; stress response.

OS Oryza sativa.

PN WO2004061122-A2.

PD 22-JUL-2004.

XX 23-DEC-2003; 2003WO-US041200.

XX 26-DEC-2002; 2002US-0436565P.

XX (SYGN) SYNGENTA PARTICIPATIONS AG.

XX Cooper B;

XX WPI; 2004-534388/51.

XX New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.

PS Claim 28; SEQ ID NO 14; 408pp; English.

XX The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related polypeptide. The
CC present sequence is published separately from the main body of the
CC specification as EPO data.

XX Sequence 379 AA;

Query Match 35.4%; Score 466.5; DB 8; Length 379;
Best Local Similarity 37.8%; Pred. No. 5.3e-35;
Matches 111; Conservative 43; Mismatches 77; Indels 63; Gaps 8;

QY 3 PPRGAAAATAALDTGVHLEASVPP-----LPEAGGNVORKGAVDP-----48

DB 37 PPSGGAQSASTSG-----GSAGSPSRSEQHVPAAGMAAGAAATPISNTFL 86

QY 49 -----DKDRKEKAAAPRITGWLREYSKIYCEKVEAKGRITYNE 88

DB 87 RLNDLIHGDDAPSSQAPTSKKKRGARAVGPKGKGLRQFSMKVCEKVESKGRITYNE 146

QY 89 VADEIYSELKSMAH-----GQGFDEKNIRRVYDAFNVLIALRVIAKEKEIRW 138

DB 147 VADELVAEFADPNNSILPPDPDPNPAQYDEKNIRRVYDALNVLNMAELISKDKKEIQW 206

QY 139 MGLSNTRYEKIKKLEVRKELV---NKIRNKALLQIEKQFDLQNIKLNRQTLSSAE 195

DB 207 KGLPR---TSINDIEDLQVELGKRIEKNYTLQELQDFVGMQKLIQNEQLVSG- 262

QY 196 NV---NGIRLPFLVKTSRKARVEIISDDSKFAHFEFNGCAPFTLHDDLILLEGV 247

DB 263 NIPSGGVALPFILVQTRPHATVEVEISDMQLVHFDNFSTPFELHDDSFVLKAM 316

RESULT 11

AAB07975

ID AAB07975 standard; protein; 299 AA.

XX AAB07975;

XX 14-NOV-2000 (first entry)

DE A corn transcription factor designated DP.

XX DP polypeptide; transcription factor; gene transcription; cell cycle;

KW DNA replication; DRTP; differentiation-regulated transcription factor 1;
KW transgenic plant; transformation efficiency.

OS Zea mays.

XX Key Location/Qualifiers

XX Misc-difference 1 /note= "any amino acid encoded by NNN"

PN WO200047614-A1.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-US003651.

XX 12-FEB-1999; 99US-0119857P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Lowe KS, Gordon-Kamm WJ, Bailey MA, Gregory CA;

XX WPI; 2000-524530/47.

XX N-PSDB; AAA59702.

XX Transgenic plants comprising an expression cassette consisting of a DP
PT nucleic acid, when expressed, the nucleic acid increases the efficiency
PT of plant cell transformation by increasing the number of dividing cells
PT in the plant.

XX Disclosure; Page 71-72; 76pp; English.

XX The present sequence represents a corn DP (not defined) polypeptide. The
CC polypeptide is a transcription factor that activates transcription of
CC numerous genes involved in DNA replication, thus playing a role in the G1
CC to S transition in the cell cycle. Do, also called DRTP (differentiation-
CC regulated transcription factor 1) has been shown to form specific
CC multiprotein complexes with the retinoblastoma susceptibility protein,
CC p107, cyclins and cdk2. The DP polynucleotide was used to produce a
CC transgenic plant. The DP nucleic acids and proteins are useful for
CC increasing transformation efficiency in plants by increasing the number
CC of dividing cells in the plant (dividing cells may be more receptive to
CC transformation)

XX Sequence 299 AA;

Query Match 35.0%; Score 461; DB 3; Length 299;
Best Local Similarity 40.7%; Pred. No. 1.2e-34;
Matches 110; Conservative 41; Mismatches 71; Indels 48; Gaps 8;

QY 16 LDLTGVHLEASVPPLEAGGNVORKG--AVDPDKRKEKAAAPRITGWLREYSKI 73

DB 11 LDING-----DDAPSSQAPTSKKKRRGTRAVGPKDKNR-----GLRQFSMK 51

QY 74 VCEKVEAKGRITYNEVADRIYSELKSMAH-----IGQGFDEKNIRRVYDAFNVL 123

DB 52 VCEKVESKGRITYNEVADLVAEFTDPNNIEAPDPDPNPAQYDEKNIRRVYDALNVL 111

QY 124 IALRVIAKEKEIRWGLSNTRYEKIKKLEVRKELV---NKIRNKALLQIEKQFDL 180

DB 112 MAMDIIISKDKKEIQWKGKLP---TSISDTEEMKTELVLGKRIEKKSAVLAQLQDQVGL 168

QY 181 QNLIKRN-OTLESSAENVNGIRLPFLVKTSRKARVEIISDDSKFAHFEFNGCAPFTLHD 239

DB 169 QNLIORNEQSYSGNTSPGGVALPFILVQTRPHATVEVEISDMQLVHFDNFSTPFELHD 228

QY 240 DLSILEGVR-----RNSIGRAGRAT 259

DB 229 DSYVLKEMRFCCREQHDSIQESISNGESS 258

RESULT 12

AAG22852

ID AAG22852 standard; protein; 277 AA.


```

PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 14-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 18-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 33.1%; Score 437; DB 3; Length 277;
Best Local Similarity 50.1%; Pred. No. 2e-32;
Matches 91; Conservative 32; Mismatches 52; Indels 6; Gaps 2;

QY 74 VCEKVEAKGRTTNEVADIEYSELKSMHIG-----QGFDEKIRRRVYDAFNVLIALRV 128
DB 3 VCEKVESKGRITTEVADIEVAFALPNNDGTSPDQQQYDEKIRRRVYDALNVLAMMDI 62
QY 129 IAKEKEIRWGLSNRYEKKLEBRKELVNKIRKALLQIEIKQFDDQLQNIKLKRNQ 188
DB 63 ISKDKKEIQWRGLPRTSLSDIELKNERLSLRNRIEKKTAYSQELBEQVGLQNLQJRN 122
QY 189 TLESSAENVN-GIRLPFLVKTSRKARVEIISDDSKFAHFFENGAPFTLHDDLSILEGV 247
DB 123 HLYSSGNAPSGGVALFPIVQTRPHATVEIVEISDMQLVHFDNFSPFELHDDNFVLKTM 182
QY 248 R 248
DB 183 K 183

```

RESULT 13

AAU72529

ID AAU72529 standard; protein; 292 AA.

XX AAU72529;

XX 26-FEB-2002 (first entry)

DE Arabidopsis cell cycle protein CCP16.

XX Cell cycle protein; CCP; cell cycle regulation; herbicide;

KW plant growth regulator; plant development; abiotic stress; biotic stress;

KW nutrient deprivation; pathogen attack; crop yield; immunogen.

XX Arabidopsis thaliana.

XX WO200185946-A2.

XX 15-NOV-2001.

XX 14-MAY-2001; 2001WO-IB001307.

XX 12-MAY-2000; 2000US-0204045P.

XX (CROP-) CROPDISEIGN NV.

PI Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX WPI; 2002-062249/08.

DR N-PSDB; AAS96319.

XX New cell cycle protein and nucleic acid molecule encoding it useful for
 PT regulating cell cycle progression in plants and for identifying
 PT modulators which are useful as herbicides or plant growth regulators.

Claim 34; Fig 16; 316pp; English.

XX The invention relates to a novel cell cycle protein (CCP) and the
 CC polynucleotides encoding them. CCP is useful for identifying a compound
 CC which modulates the activity of the polypeptide and which binds to the
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP
 CC nucleic acid and polypeptide molecules are useful as modulating agents in
 CC regulating cell cycle progression in plants. CCP is useful to treat
 CC disorders characterised by insufficient or excessive production of CCP
 CC protein or production of CCP protein forms which have decreased or
 CC aberrant activity. Compounds that bind to or modulate the activity of CCP
 CC polypeptide are useful as herbicides or plant growth regulators. The
 CC polynucleotide is useful for modifying cell fate, plant development,
 CC plant morphology, biochemistry and/or physiology, the length of the G1,
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
 CC stimulation or enhancement of cell division, DNA replication, seed set,
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
 CC initiation and/or development, nodule function, dwarfism in plants, senescence,
 CC tolerance or resistance to stress. CCP, the polynucleotide
 CC and the anti-CCP antibody are useful in agriculture to modulate the
 CC protein levels or activity of a protein involved in the cell cycle due to
 CC environmental conditions, including abiotic stress such as cold, nutrient
 CC deprivation, heat, drought, salt stress, or biotic stress such as
 CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate
 CC plant architecture, plant quality traits, plant reproduction and seed
 CC development, endoreduplication in storage cells, storage tissues and/or
 CC storage organs of plants or its parts. CCP is useful as an immunogen to
 CC generate antibodies. CCP protein is useful to screen for naturally
 CC occurring CCP substrates. The polynucleotide is useful for expressing CCP
 CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to
 CC modulate CCP activity. The present sequence represents a CCP protein of
 CC the invention

```
SQ Sequence 292 AA;
Query Match 32.5%; Score 428.5; DB 5; Length 292;
Best Local Similarity 41.3%; Pred. NO. 1.4e-31;
Matches 88; Conservative 49; Mismatches 69; Indels 7; Gaps 2;

Qy 40 VORKGAVDPDKRKKKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADSIYSELKS 99
Db 27 VRRKLIVDDSEIGSEKKQSRTSGGGLRQFVWVCQKLEAKITTYKEVADSIISDFAT 86
Qy 100 MAHIGQ-----GFDEKNIRRRVYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKKLE 153
Db 87 IKQNAEKPLNENEYNEKNIRRRVYDALNVFMDLIARDKKEIRWGLPTCKKQVDEVK 146
Qy 154 EVRKELVKNIRNKALLQIEKQFDLQNIKLNRQNTLESSAEN-VNGIRLPFVLVTSRK 212
Db 147 MDRNKMVSSVQKKAAPFLKELREKVSLESLSMRNQEMVVKTOGPARGFTLPFILLETNPH 206
Qy 213 ARVEIEISDDSKFAHPEFNGAPFTLHDDLILE 245
Db 207 AVVEIEISEDMQLVHLDFNSTPFSVHDDAYILK 239

RESULT 14
ABP96848
ID ABP96848 standard; protein; 292 AA.
XX AC ABP96848;
XX DT 10-JUN-2003 (first entry)
XX DE Arabidopsis thaliana DP protein SEQ ID NO.4.
XX KW Arabidopsis thaliana; modulation; endoreduplication; plant;
XX KW transgenic plant; cell cycle; E2F; DP.
XX OS Arabidopsis thaliana.
XX PN WO2003018818-A2.
XX PD 06-MAR-2003.
XX PF 26-AUG-2002; 2002WO-EP009504.
XX PR 24-AUG-2001; 2001US-00938342.
XX PA (CROP-) CROPDESIGN NV.
XX PI De Veylder L, Inze D, Mironov V, Segers G;
XX WPI; 2003-300735/29.
XX DR N-PSDB; ACC45100.
XX PT Modulating endoreduplication in a plant, or a part of the plant,
XX PT comprises modifying the expression or activity of an E2Fa and DPa gene or
XX PT polypeptide.
XX FS Disclosure; Page 26-27; 34pp; English.
XX CC The present invention describes a method for modulating endoreduplication
XX CC in a plant, or a part of the plant. The method comprises modifying the
XX CC expression or activity of an E2F and DP gene or polypeptide. Also
XX CC described: (1) a transgenic plant cell overexpressing an E2F or E2Fa
XX CC gene, or an E2F and DP gene or an E2Fa and DP gene, where the E2F or
XX CC E2Fa gene or the DP or DPa gene is under the control of a tissue or cell-
XX CC type specific promoter; (2) a transgenic plant, or a part of the plant,
XX CC comprising the cells described above; (3) a progeny of the plant
XX CC described above; and (4) a plant material obtained from the plant
XX CC described above. The method is useful in modulating plant cell cycle
XX CC proteins by modifying the expression or activity of an E2F and DP gene or
XX CC polypeptide. The present sequence represents Arabidopsis thaliana DP,
XX CC which is used in the exemplification of the present invention
```

```
SQ Sequence 292 AA;
Query Match 32.5%; Score 428.5; DB 6; Length 292;
Best Local Similarity 41.3%; Pred. NO. 1.4e-31;
Matches 88; Conservative 49; Mismatches 69; Indels 7; Gaps 2;

Qy 40 VORKGAVDPDKRKKKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADSIYSELKS 99
Db 27 VRRKLIVDDSEIGSEKKQSRTSGGGLRQFVWVCQKLEAKITTYKEVADSIISDFAT 86
Qy 100 MAHIGQ-----GFDEKNIRRRVYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKKLE 153
Db 87 IKQNAEKPLNENEYNEKNIRRRVYDALNVFMDLIARDKKEIRWGLPTCKKQVDEVK 146
Qy 154 EVRKELVKNIRNKALLQIEKQFDLQNIKLNRQNTLESSAEN-VNGIRLPFVLVTSRK 212
Db 147 MDRNKMVSSVQKKAAPFLKELREKVSLESLSMRNQEMVVKTOGPARGFTLPFILLETNPH 206
Qy 213 ARVEIEISDDSKFAHPEFNGAPFTLHDDLILE 245
Db 207 AVVEIEISEDMQLVHLDFNSTPFSVHDDAYILK 239

RESULT 15
ADH59560
ID ADH59560 standard; protein; 292 AA.
XX AC ADH59560;
XX DT 25-MAR-2004 (first entry)
XX DE Dimerisation partner.
XX KW E2F transcription factor; Gene therapy; stress resistance;
XX KW differentiation signal.
XX OS Arabidopsis thaliana.
XX PN WO2003025185-A1.
XX PD 27-MAR-2003.
XX PF 12-SEP-2002; 2002WO-EP010236.
XX PR 14-SEP-2001; 2001EP-00870198.
XX PA (CROP-) CROPDESIGN NV.
XX PI Beekman T, De Veylder L, Inze D, Mironov V, Broekaert W;
XX PI Dillen W, Frankard V;
XX WPI; 2003-371815/35.
XX DR N-PSDB; ADH59559.
XX PT Modifying cell number, architecture and yield of plants comprises
XX PT modulating the expression and/or the activity of members of the plant E2F
XX PT transcription factor family.
XX PS Claim 18; SEQ ID NO 4; 41pp; English.
XX CC The present invention relates to increasing the cell number of specific
XX CC cell types, specific tissues or specific organs in a plant comprises
XX CC modulating the expression and/or the activity in the specific cell types,
XX CC specific tissues or specific organs of the plant E2F transcription
XX CC factor. The E2F transcription factor or its homologue, derivative or
XX CC fragment, is useful for prolonging the period of cell division in certain
XX CC cells and tissues, for increasing the size of cotyledons, for enhancing
XX CC cell proliferation after seed germination, for enhancing stress
XX CC resistance of seedlings, for obtaining seedlings with enhanced vigor, for
XX CC obtaining plants that have increased growth or number of organs, for
XX CC obtaining plants having more cells in a particular tissue, for obtaining
XX CC plants having an increased size of organs or having increased yield, for
XX CC stimulating differentiated cells to re-enter the cell cycle, for
```

CC overriding the cell differentiation signals, or for altering cell shape.
CC The present sequence represents the dimerisation partner.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

SQ Sequence 292 AA;

Query Match	32.58	Score 428.5	DB 7	Length 292
Best Local Similarity	41.3	Pred. No. 1.4e-31		
Matches	88	Conservative 49	Mismatches 69	Indels 7
Gaps	2			
Qy	40	VORKGAVDPDKRAKKEKAAAPRITGMGLRYSKIVCEKVEAKGRTTIVNEVADEIYSLKS	99	
Db	27	VRKKLI VDDSDSEIGSEKGGQSRSTGGGLRQFSVMVCQKLEAKKITTITTYKEVADEIISDFAT	86	
Qy	100	MAHIGQ-----GFDEKNIRRRYDAPNVLIALRVIAKEKEIRWGLSNRYRETKKLE	153	
Db	87	IKQNAEKPLNENYEKNIRRRYDAPNLFMDIIARDKKEIRWGLPLTICKCKDVEEVK	146	
Qy	154	EVREKLNNKIRNKALIQETEKQFDLQNTLRNQTLESSAEN-VNGIRLPFVLVKTSRK	212	
Db	147	MDRNKVMSSVQKAAFLKELREKVSSLESLSMRNQEMVMTQGPAGFTLPFLILEINPH	206	
Qy	213	ARVEIIEISDSKFAHPEFNGAPTLHDDLSILE	245	
Db	207	AVVEIIEISDMQLVHLDFNSTPFSVHDDAYILK	239	

Search completed: February 4, 2005, 20:30:04
Job time : 168 secs

This Page Blank (uspto)

A;Residues: 1-377 <DYN>
A;Cross-references: UNIPROT:Q24318; GB:X79708; NID:G516866; PID:G516867
C;Genetics:
A;Gene: FlyBase:Dp
A;Cross-references: FlyBase:FBgn0011763
C;Superfamily: transcription factor DP

Query Match 28.5%; Score 375.5; DB 2; Length 377;
Best Local Similarity 39.1%; Pred. No. 1.8e-17;
Matches 90; Conservative 49; Mismatches 68; Indels 23; Gaps 8;

QY 21 VHILEASSVPLPRAGGNVAVKRGAVDPDKDRKKKAAAPRITGWGLREYSKIVCEKVEA 80
DB 71 LHAIQSNLSHM-SASSSSVORK-----RDKA-----GKGLRHFSMKVCEKVEE 115
QY 81 KGRITNEVDAEYIS-ELKSMHAHQGGFDEKNIRRRVYDAFNVLIALRVIATAKEKEIRWM 139
DB 116 KGKTTYNEVADDLVSEEMKNAY-DNDCQKNIRRRVYDALNVLMAINVISKDKKEIRWI 174
QY 140 GLSNRYVEKIKLBEVRKELVNIRKALKQIEKQFDLQNIKURNQTLSS-----AE 195
DB 175 GLPANSTETFLALSEENCQRREIRKQKQEMLRIMQHVAFGLVERNKRNESQGVVPS 234
QY 196 NVNGIRLPFVLVKTSSRKARVIEISDDSKFAHFEENGAPFTLHDDL SILE 245
DB 235 NAS-IQLPFIIVNTHKSTKNCVNTDKSEYIFKFD-KTFEMHDDIEVLK 282

RESULT 7
A48585
transcription factor DP-1 - human
C;Species: Homo sapiens (man)
C;Date: 11-Nov-1994 #sequence_revision 11-Nov-1994 #text_change 09-Jul-2004
C;Accession: A48585
R;Helin, K.; Wu, C.L.; Fattaey, A.R.; Lees, J.A.; Dynlacht, B.D.; Ngwu, C.; Harlow, E.
Genes Dev. 7, 1850-1861, 1993
A;Title: Heterodimerization of the transcription factors E2F-1 and DP-1 leads to cooperation in cell cycle regulation
A;Reference number: A48585; MUID:94010284; PMID:8405995
A;Accession: A48585
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-410 <HEL>
A;Cross-references: UNIPROT:Q14186; GB:L23959; NID:G414316; PIDN:AAA58440.1; PID:G414317
C;Superfamily: transcription factor DP
C;Keywords: DNA binding; transcription factor

Query Match 28.4%; Score 374.5; DB 2; Length 410;
Best Local Similarity 42.0%; Pred. No. 2.3e-17;
Matches 86; Conservative 42; Mismatches 60; Indels 17; Gaps 7;

QY 50 KDRKKEAAAPRITGWGLREYSKIVCEKVEAKGRITTYNEVADSIYSELKSM-AHI---GQ 105
DB 105 RNRKGEK-----NGKGLRHFSMKVCEKQVRKGTTSYNEVADELVAEFSADNHLPNES 158
QY 106 GFDEKNIRRRVYDAFNVLIALRVIATAKEKEIRWMGLSNRYVEKIKLBEVRKELVNIRN 165
DB 159 AYDQKNIRRRVYDALNVLMAINVISKDKKEIRWI-----GKGLRHFSMKVCEKVEE 218
QY 166 KXALLOIEKQFDLQNIKURNQTLSSAENV-----NGIRLPFVLVKTSSRKARVIEISD 221
DB 219 KQSLOQLIQQIAFKNIVQRNRAHQQAASRPPPNVSIHLPFIIVNTSKKTVIDCSISN 278
QY 222 DSKFAH-FEENGAPFTLHDDL SILE 245
DB 279 D-KPEYLFNFNT-FEIHDDIEVLK 301

RESULT 8
T25207
hypothetical protein T23G7.1 - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C;Accession: T25207

R;Barlow, K.
Submitted to the EMBL Data Library, December 1995
A;Reference number: Z19995
A;Accession: T25207
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-598 <WIL>
A;Cross-references: UNIPROT:Q22703; EMBL:Z68319; PIDN:CAA92699.1; GSPDB:GN00020; CESP:T2
A;Experimental source: clone T23G7
C;Genetics:
A;Gene: CESP:T23G7.1
A;Map position: 2
A;Introns: 3/3; 115/3; 204/2; 227/2; 260/1; 298/1; 363/1; 424/3; 478/3; 565/1

Query Match 26.8%; Score 354; DB 2; Length 598;
Best Local Similarity 41.3%; Pred. No. 7.9e-16;
Matches 78; Conservative 34; Mismatches 67; Indels 10; Gaps 3;

66 GLREYSKIVCEKVEAKGRITTYNSVADEIYSE-----LKSMHIGQGFDEKNIRRVYDA 119
Db GLRHFTKVCVKKEKGLTNYNSVADLVADYFQNNLIKQIDVVQKQYDMKNIRRVYDA 131

120 FNVLIARLVAKEKKIRTMGLSNRYEKIKLEEVKELVNKRNKALLQIEKQFDD 179
Db LNVLLAMNITTKKQIRWIGLSPASQELSRLEEKSRREASISKKQALEEWLIQVS 191

180 LQNIKLNRNLTLE---SSAENVNGIRLPVLVKTSRKARVEIBISDDSKFAHFFNGAPFT 236
Db YKLVNERNRKHNGKRPENDTVLHLPFLINTKEANVECSVSDKSEFLFSFD-KKFE 250

237 LHDDLSTLE 245
251 IHDDFEILK 259

RESULT 9
JC4929
transcription factor E2F1 - human
N;Alternate names: Ap12 protein; retinoblastoma-associated protein; retinoblastoma-bind
C;Species: Homo sapiens (man)
C;Date: 22-Oct-1996 #sequence revision 01-Nov-1996 #text_change 09-Jul-2004
C;Accession: JC4929; A45032; A42998; A42997; I54091
R;Neuman, E.; Sellers, W.R.; McNeil, J.A.; Lawrence, J.B.; Kaelin Jr., W.G.
Gene 173, 163-169, 1996
A;Title: Structure and partial genomic sequence of the human E2F1 gene.
A;Reference number: JC4929; MUID:97082961; PMID:8964493
A;Accession: A42998
A;Status: preliminary; nucleic acid sequence not shown
A;Molecule type: DNA
A;Residues: 1-437 <NEU>
A;Cross-references: UNIPROT:Q01094; GB:U47675; NID:g1594281; GB:U47676; NID:g1594282; GB
A;Experimental source: placenta
R;Shan, B.; Zhu, X.; Chen, P.L.; Durfee, T.; Yang, Y.; Sharp, D.; Lee, W.H.
Mol. Cell. Biol. 12, 5620-5631, 1992
A;Title: Molecular cloning of cellular genes encoding retinoblastoma-associated proteins
A;Reference number: A45032; MUID:93078763; PMID:1448092
A;Accession: A45032
A;Molecule type: mRNA
A;Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>
A;Cross-references: GB:S49592; NID:g260573; PIDN:AA324289.1; PID:g260574
A;Note: sequence extracted from NCBI backbone (NCBIP:119095)
A;Note: the authors are uncertain whether Met-1 is the initiator or whether translation
R;Helin, K.; Lees, J.A.; Vidal, M.; Dyson, N.; Harlow, E.; Fattaey, A.
Cell 70, 337-350, 1992
A;Title: A cDNA encoding a PRB-binding protein with properties of the transcription fact
A;Reference number: A42998; MUID:92346720; PMID:1638634
A;Accession: A42998
A;Molecule type: mRNA
A;Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <HEB>
A;Cross-references: GB:N96577; NID:g181917; PIDN:AA35782.1; PID:g181918
A;Experimental source: Nalm 6 pre-B leukemia cell line
A;Note: sequence extracted from NCBI backbone (NCBIN:110015, NCBIP:110016)
R;Kaelin Jr., W.G.; Krek, W.; Sellers, W.R.; DeCaprio, J.A.; Ajchenbaum, F.; Fuchs, C.S.

Cell 70, 351-364, 1992
A;Title: Expression cloning of a cDNA encoding a retinoblastoma-binding protein with E2F
A;Reference number: A42997; MUID:92346721; PMID:1638635
A;Accession: A42997
A;Molecule type: mRNA
A;Residues: 1-312, 'S', 314-321, 'N', 323-328, 'T', 330-437 <KAE>
A;Cross-references: GB:U13851; NID:g595713
A;Experimental source: Akata cells; expression vector pGEX-2TK
A;Note: sequence extracted from NCBI backbone (NCBIN:110018, NCBIP:110019)
R;Johnson, D.G.; Ohtani, K.; Nevins, J.R.
Genes Dev. 8, 1514-1525, 1994
A;Title: Autoregulatory control of E2F1 expression in response to positive and negative
A;Reference number: I54091; MUID:95047311; PMID:7958836
A;Accession: I54091
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-88, 'R', 'T', '122-123, 'TPR', '127, 'QRR', '297-299, 'PRR', '308-309, 'RA', '312, 'C', '298-
A;Cross-references: GB:S74230; NID:g712816; PIDN:AAD14150.1; PID:g4261850
C;Genetics:
A;Gene: GDB:E2F1
A;Cross-references: GDB:134661; OMIM:189971
A;Map position: 20q11-20q11
A;Introns: 87/3 118/1 191/2 242/2 280/3 356/1
C;Keywords: DNA binding; transcription factor
F;67-108/Region: cyclin box #status predicted
F;118-190/Domain: DNA binding #status predicted <DNA>
F;191-241/Region: 7-residue repeats

Query Match 9.7%; Score 128.5; DB 2; Length 437;
Best Local Similarity 24.3%; Pred. No. 0.32;
Matches 64; Conservative 33; Mismatches 97; Indels 69; Gaps 10;

QY 2 APPRGAAAATAALDLTGWHILEAS-----SVPPLEAGGNVQKGVDPD- 49
Db APAGGCPAPALEALLGAGALLDSSQIVITISAAQDASAPPAPT--GPAAPAGPCDDPL 63

QY 50 -----KDRKKEKAAAPR-----ITGWLREYSKIVCEK 77

Db 64 LIFATPQAPRPTSPAPRALGPPVKKRDLDTDHOYLAESSGPARGRHPKGVKSPG 123

QY 78 VBAGRTTYNEVADEIYSELKSMHIGQGFDEKN-----IRREYDAFNVLIALRVI 129

Db 124 EKSRVETSLNLTTRF---LELLSHSADGVDLNAAEVLKVKGRYIDITNVLGIQLI 190

QY 130 AKX-KKEIRWMLGSLNRYEIKKLEEVKELVNKRNKALLQIEKQFDDLQNIKLNRQ 198

Db 181 AKKKNHIOWLQ-SHTTVGVGRLEGLTQDL-----ROLQESQQLDHLNICTTQL 231

QY 189 TLESSAENVNGIRLPVLVKTSR 211

Db 232 RLLS--EDTDSORLAYVTQDLR 252

RESULT 10
T31602
hypothetical protein Y48C3A.t - Caenorhabditis elegans
C;Species: Caenorhabditis elegans
C;Date: 29-Oct-1999 #sequence revision 29-Oct-1999 #text_change 29-Oct-1999
C;Accession: T31602
R;Wallis, J.
submitted to the EMBL Data Library, September 1999
A;Reference number: Z21046
A;Accession: T31602
A;Status: preliminary; translated from GB/EMBL/DBDJ
A;Molecule type: DNA
A;Residues: 1-470 <WIL>
A;Cross-references: EMBL:AL117203; NID:e1549827; PIDN:CAB55117.1; CESP:Y48C3A.t
A;Experimental source: clone Y48C3A
C;Genetics:
A;Gene: CESP:Y48C3A.t
A;Introns: 52/3; 156/2; 206/2; 237/3; 332/1; 386/1

Query Match 9.1%; Score 120.5; DB 2; Length 470;

C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004
C/Accession: T27066
R/McMurray, A.
submitted to the EMBL Data Library, January 1998
A/Reference number: Z20305
A/Acession: T27066
A/Status: preliminary; translated from GB/EMBL/DDBJ
A/Molecule type: DNA
A/Residues: 1-717 <WIL>
A/Cross-references: UNIPROT:O45958; EMBL:AL021493; PIDN:CAA16394.1; GSPDB:GN00023; CESP:
A/Experimental source: clone Y51A2B
C/Genetics:
A/Gene: CESP:Y51A2B.6a
A/Map position: 5
A/Introns: 8/1; 33/1
C/Superfamily: Caenorhabditis elegans hypothetical protein Y116A8C.33

Query Match 8.5%; Score 112.5; DB 2; Length 717;
Best Local Similarity 26.0%; Pred. No.6.6;
Matches 57; Conservative 35; Mismatches 54; Indels 73; Gaps 14;

	Qy	87	NVADIEIY-SELKS-MAHIGQGDEKNIIRRVYDAFNVLIALRLVIAGEKK-----	134
	Dy	368	NSPVDENFESOFKTMLNVG-----KVQDLFNVLVDLSSTENKFKLKRKMVLLS	417
	Qy	135	-----ETRWGLSNY-----RYEK-IKKLEEVRELVNKRN	165
	Dy	418	PINLDKHLIRVKVEGLEEQLQTSLYYEKLPFPQPYRKYSKYLKLLKLNGRKEFAKDND	477
	Qy	166	KKAL-- --LOIEKGFDLDLNT---- -KLNRQTLESSAEENVGIRLPFVL---VKTSRKA	214
	Dy	478	ELSIVIKCLQQLKQSDFHFQKIQVIQWRNLDD-KDSQNIIQGI---PSVVSEFLKNLAQR	534
	Qy	215	VFEISD ----DKFAHFEPNGAPFTTLHDLL-SILEGYR	248
	Dy	535	----EISDKMKENSNNKTFVMNALSFNKSDSIKSIVSSVQ	570

RESULT 13
T27067
hypothetical protein Y51A2B.6b - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 09-Jul-2004

```

C;Accession: 127067
R;McMurray, A.
submitted to the EMBL Data Library, January 1998
A;Reference number: Z20305
A;Accession: T27067
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: DNA
A;Residues: 1-717 <WIL>
A;Cross-references: UNIPROT:O45959; EMBL:AL021493; PIDN:CAA16395.1; GSPDB:GN00023; CESP:Y
A;Experimental source: clone Y51A2B
C;Genetics:
A;Gene: CESP:Y51A2B.6b
A;Map position: 5
A;Introns: 8/1; 33/1
C;Superfamily: Caenorhabditis elegans hypothetical protein Y116A8C.33

Query Match      8.5%; Score 112.5; DB 2; Length 717;
Best Local Similarity 26.08; Pred. No. 6.6;
Matches 57; Conservative 35; Mismatches 54; Indels 73; Gaps 14;

QY      87 NEVADEIY-SELKS-MAHIGQGDEKNIRRVYDAFNVLIALRVIAKEKK----- 134
      |  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      368. NSPDEKFESQFKTLMNVG-----KVQLFNVLVDLSSIEKFKKLRKVMLLS 417
      |  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY      135 -----EIRWMGLSNY-----RYEK- IKKLBEVRKELVNKIRN 165
      |  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
Db      418 PINLDKHLIRKVEGIIEELQQLTSLYYEKLPLFPQYRKYKLLKKLINGRKEFADKVDN 477
      |  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

QY      166 KKAL----LOBIEKQFDPLQNI-----KLRNOTLESSAENVNGIRLPFVL----VKTSRKAR 214
      |  ||  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |

```

```

Db      418 PINLDHLIRKVEGILEELQQLTSLYYEKLLPFQPVRYKKLKLKLNCRKEFADVDN 477
Qy      166 KKAL---LQIEKQFDLQNI-----KLRNQTLESSAENVNGIRLPFVL---VKTSRKAR 214

```


GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2005, 20:06:49 ; Search time 170 Seconds
(without alignments)
786.192 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGAAAAATAALDLTG.....SILEGVRNSIGRAGRATLH 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt 03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1304	98.9	261	2 Q9FET1	Q9fet1 triticum sp
2	890.5	67.5	294	2 Q84VF4	Q84vf4 oryza sativ
3	838	63.5	289	2 Q8S182	Q8s182 oryza sativ
4	492	37.3	385	2 Q8LDG3	Q8ldg3 arabidopsis
5	492	37.3	385	2 Q9FNY2	Q9fny2 arabidopsis
6	476.5	36.1	353	2 Q7XZ14	Q7xz14 populus tre
7	471.5	35.7	346	2 Q84VD5	Q84vd5 oryza sativ
8	466.5	35.4	379	2 Q84VA0	Q84va0 oryza sativ
9	456	34.6	413	2 Q9LZ27	Q9lze7 arabidopsis
10	428.5	32.5	288	2 Q9LZ55	Q9lze5 arabidopsis
11	428.5	32.5	292	2 Q9FNY3	Q9fny3 arabidopsis
12	399.5	30.3	388	2 Q6R754	Q6r754 homo sapien
13	394	29.9	359	2 Q9UG28	Q9ug28 homo sapien
14	394	29.9	446	1 TDP2 HUMAN	Tdp2 human
15	393	29.8	395	2 Q8BHD2	Q8bhd2 m mus muscu
16	388	29.4	385	2 Q8C537	Q8c537 mus musculu
17	382	29.0	409	2 Q63ZK9	Q63zr9 xenopus lae
18	382	29.0	412	2 Q66K12	Q66k12 xenopus lae
19	378.5	28.7	396	2 Q9D297	Q9d297 mus musculu
20	378.5	28.7	410	1 TDP1 MOUSE	Tdp1 mouse
21	375.5	28.5	377	1 TDP1 DROME	Tdp1 drosophila
22	375.5	28.5	441	2 Q8ML56	Q8ml56 drosophila
23	375.5	28.5	441	2 Q8PE60	Q8pe60 drosophila
24	375.5	28.5	445	2 Q9V6M0	Q9v6m0 drosophila
25	374.5	28.4	410	1 TDP1 HUMAN	Tdp1 human
26	373	28.3	409	2 Q6TNQ1	Q6tnq1 brachydanio
27	372.5	28.2	318	2 Q7XE27	Q7xe27 oryza sativ
28	372	28.2	409	2 Q803N1	Q803n1 brachydanio
29	364.5	27.6	363	2 Q44080	Q44080 drosophila
30	356	27.0	506	2 Q7PFI9	Q7pfi9 anopheles g
31	356	27.0	584	2 Q7PSH2	Q7psh2 anopheles g

32 355 26.9 310 2 Q8C8M7
33 354 26.8 598 1 TDPI CAEEL
34 351.5 26.6 381 2 Q9GT27
35 349.5 26.5 290 2 Q9CY27
36 345.5 26.5 386 2 Q6PBR6
37 323 24.5 446 1 TDP2 MOUSE
38 302 22.9 345 2 Q9NZ54
39 234.5 17.8 657 2 Q7KWP1
40 165 12.5 212 2 Q8SR59
41 128.5 9.7 437 1 E2F1 HUMAN
42 120.5 9.1 287 2 Q9BIE8
43 120.5 9.1 412 2 Q9U289
44 117 8.9 425 2 Q654W1
45 116.5 8.8 864 1 RA50_SULSO

ALIGNMENTS

RESULT 1

Q9FET1 PRELIMINARY; PRT; 261 AA.
AC Q9FET1; (TREMBlrel. 16, Created)
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE DP protein.
OS Triticum sp.
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Pooideae;
OC Triticeae; Triticum.
OX NCBI_TaxID=4569;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20362804; PubMed=11108846; DOI=10.1016/S0014-5793(00)02239-0;
RA Ramirez-Parra E., Gutierrez C.;
RT "Characterization of wheat DP, a heterodimerization partner of the
RT plant E2F transcription factor which stimulates E2F2DNA binding.";
RL FEBS Lett. 486:73-78(2000).
CC -!- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -!- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AJ271917; CAC19034.1; -
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009058; WING_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 261 AA; 29262 MW; 5008D8F193163A28 CRC64;

Query Match 98.9%; Score 1304; DB 2; Length 261;
Best Local Similarity 98.9%; Pred. No. 9.5e-77;
Matches 258; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 MAPPRGAAAAATAALDLTGVIHLEASSVPPIPEAGNVAQRKGAVDPKDRKKEKAAAP 60
DB 1 MAPPRGAAAAATAALDLTGVIHLEASSVPPIPERGNAVQRKGAVDPKDRKKEKAAAP 60
QY 61 RITGWLREYSKIVCEKAGRTTNEVADEIYSELKSMHITGQGFDEKNTRRRYVDAF 120
DB 61 RITGWLREYSKIVCEKAGRTTNEVADEIYSELKSMHITGQGFDEKNTRRRYVDAF 120
QY 121 NVLIARLVIATKKEKIRWMLNSRYEYKIKLEEVRKELVKNIRKALLQIEKEQFDDL 180
DB 121 NVLIARLVIATKKEKIRWMLNSRYEYKIKLEEVRKELVKNIRKALLQIEKEQFDDL 180
QY 181 QNKLNRNQTLESSAENVNGIRLPFLVKTSRKARVEIEISDDSKFAHFENGFAPFTLHDD 240
DB 181 QNKLNRNQTLESSAENVNGIRLPFLVKTSRKARVEIEISDDSKFAHFENGFAPFTLHDD 240

```

QY 241 LSILEGVRRNSIGRAGRATLH 261
Db 241 LSILEGVRRNSIGRAGRATLH 261

RESULT 2
Q84VF4 PRELIMINARY; PRT; 294 AA.
ID Q84VF4;
AC Q84VF4;
DT 01-JUN-2003 (TREMBlrel. 24, Created)
DT 01-JUN-2003 (TREMBlrel. 24, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE DP protein.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=23111120; PubMed=14750518;
RX DOI=10.1023/B:PLAN.000007001.30865.0f;
RA Cooper B., Hutchison D., Park S., Guimil S., Luginbuhl P., Ellero C.,
RA Goff S.A., Glazebrook J.
RT Identification of rice (Oryza sativa) proteins linked to the cyclin-
mediated regulation of the cell cycle."
RL Plant Mol. Biol. 53:273-279(2003).
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -|- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AV224529; AAO72649.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q84VF4; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 294 AA; 32859 MW; D1A009564782671 CRC64;

Query Match 67.5%; Score 890.5; DB 2; Length 294;
Best Local Similarity 63.1%; Pred. No. 6.5e-50;
Matches 185; Conservative 31; Mismatches 36; Indels 41; Gaps 4;

QY 1 MAPPGGAAAAATAALDLTGVLHLEASSVPPLPE----- 34
Db 1 MAPPGGAAAAAASAPGLANLLIREGAGLPSRPERYPFRPCTSDSFAPISREGDDIPPO 60

QY 35 -----AGGNAVQRK--GAVDPDKRKKKAAAPRITGWLREYSKIVCKVEAKGR 83
Db 61 KKSLSLRSGGGNAABREEGA---NRNGKKEKTGAQRITGWLREFSKIVSKVEAKGR 117

QY 84 TTNEVADEIYSELKSWAHIGQGFDEKNIRRRYDAFNVLIAIRVIAKEKEIRNWGLSN 143
Db 118 TTNEVADEIFAELKSTITQNGLEFDEKNIRRRYDAFNVLIAIRVIAKDKKEIKWGLTN 177

QY 144 YRYEKIKLEVRKELVNKIRNKALQIEIKQFDLQNTKLNRQTLSSAENVNGIRLP 203
Db 178 YRYEKIQKLEVEVKELITRIKNKKLQIEIKQFDLQNTLNRQASQRPASVNGILLP 237

QY 204 FVLVKTSRKARVEIISDDSKFAHFENGAPFTLHDDLSTILEGVRRNSIGRAG 256
Db 238 FLLIKTSRKARVEIISDSKFAHFENGAPFTMHDDVDSILEAIRRNK-CRAG 289

RESULT 3
Q8S182 PRELIMINARY; PRT; 289 AA.
ID Q8S182
AC Q8S182;

```

```

DT 01-JUN-2002 (TREMBlrel. 21, Created)
DT 01-JUN-2002 (TREMBlrel. 21, Last sequence update)
DT 01-MAR-2004 (TREMBlrel. 26, Last annotation update)
DE Putative DP protein.
GN Name=B1144G04.21;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=12447438; DOI=10.1038/nature01184;
RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
RA Wu J., Nilimura Y., Cheng Z., Hagamura Y., Antonio B.A., Kananori H.,
RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
RA Hijishita S., Honda M., Ichikawa Y., Idenuma A., Iijima M., Ikeda M.,
RA Ikeno M., Itoh S., Itoh Y., Itoh T., Iwabuchi A., Kamiya K.,
RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai, K., Shibata M.,
RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
RA Zhong H., Iwana H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
RA Yano M., Jiang J., Gojobori T.
RT The genome sequence and structure of rice chromosome 1."
RL Nature 420:312-316(2002).
CC -|- SUBCELLULAR LOCATION: Nuclear (By similarity).
CC -|- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AP003335; BAB90030.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q8S182; -.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 289 AA; 32157 MW; B911BEC6E3FFDAA CRC64;

Query Match 63.5%; Score 838; DB 2; Length 289;
Best Local Similarity 60.9%; Pred. No. 1.6e-46;
Matches 179; Conservative 31; Mismatches 36; Indels 48; Gaps 6;

QY 1 MAPPGGAAAAATAALDLTGVLHLEASSVPPLPE-----AGGNAVQR 42
Db 1 MAPPGGAAAAAASAPGLANLLIREGAGLPSRPEREGDDIPPOKKSLSLRSGGGNAABR 60

QY 43 K--GAVDPDKRKKKAAAPRITGWLREYSKIVCKVEAKGRTTNE----- 88
Db 61 EREGA---NRNGKKEKTGAQRITGWL-----LSKKVEAKGRTTNEIMVQTSNDEVYT 111

QY 89 -----VADEIYSELKSWAHIGQGFDEKNIRRRYDAFNVLIAIRVIAKEKEIRNWGLS 142
Db 112 SSGELIVADEIFAELKSTITQNGLEFDEKNIRRRYDAFNVLIAIRVIAKDKKEIKWGLT 171

QY 143 NYRYEKIKLEEVKELVNKIRNKALQIEIKQFDLQNTKLNRQTLSSAENVNGIRL 202
Db 172 NYRYEKIQKLEVEVKELITRIKNKKLQIEIKQFDLQNTLNRQASQRPASVNGILL 231

QY 203 FVLVKTSRKARVEIISDDSKFAHFENGAPFTLHDDLSTILEGVRRNSIGRAG 256
Db 232 PFLIKTSRKARVEIISDSKFAHFENGAPFTMHDDVDSILEAIRRNK-CRAG 284

RESULT 4
Q8LDG3 PRELIMINARY; PRT; 385 AA.
ID Q8LDG3
AC Q8LDG3;

```

RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=20562805; PubMed=11108847; DOI=10.1016/S0014-5793(00)02238-9;
RY	Magyar Z., Acanassova A., de Veylder L., Rombauts S., Inze D.;
RT	"Characterization of two distinct DP-related genes from Arabidopsis
RV	thaliana.";
RL	FBS Lett. 486:79-87(2000).
CC	- - SUBCELLULAR LOCATION: Nuclear (By similarity).
CC	- - SIMILARITY: Belongs to the E2F/DP family.
DR	EMBL; AJ294532; CAC15484.1; -.
DR	HSSP; Q14188; 1CF7.
DR	GO; GO:0005634; C:nucleus; IEA.
DR	GO; GO:0005667; C:transcription factor complex; IEA.
DR	GO; GO:0003700; F:transcription factor activity; IEA.
DR	GO; GO:0000074; P:regulation of cell cycle; IEA.
DR	GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR	InterPro; IPR003316; E2F TDP.
DR	InterPro; IPR003058; Wing_hlx_DNA_bnd.
DR	Pfam; PF02319; E2F_TDP; 1.
KW	DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ	SEQUENCE 385 AA; 42755 MW; 5DDB4ACA04C52AF8 CRC64;
	Query Match 37.3%; Score 492; DB 2; Length 385;
	Best Local Similarity 47.7%; Pred. No. 5.6e-24;
	Matches 104; Conservative 40; Mismatches 68; Indels 6; Gaps 2
Qy	37 GNAYQRKGVADPDKRKKKAAPRITGWLREYSKIYCEKVAKGRTTYNEVADEIYSE 96 : : : : : : : :
Dd	74 GDAGSQSGASGVKKKGQRAGDPDKTGRGLRQFSMKVCKEVSKGRTTYNEVADELVAE 133 : : : : : : : :
Qy	97 LKSWAHGT-----QGFEKNIRRRYDAFNVLIALRVIAEKKEIKRMWGLSNYRYEKIKK 151 : : : : : : : :
Dd	134 FALPNNDGTSPQQOYEQKRIIRRVYDALNLVMAMDIISKOKKETQWRGLPTSLSDIEE 193 : : : : : : : :
Qy	152 LEFYRKELVNKRNKALLQBIEKFQDDLONIKLRNQTLSSAENVN-GIRLPFLVLVKTS 210 : : : : : : : :
Dd	194 LKNERLSLNRIEKTAYSQLEEQYVGQLQNLIORNEHLYSSGNAPS GGVALPFILVQTR 253 : : : : : : : :
Qy	211 RKARVEITEISDSFAHFENFGAPFTLHDLSILEGVR 248 : : : : : : : :
Dd	254 PHATVEISEDMQLVFDFNSTFPDLHDDNFVLTKMK 291 : : : : : : : :
RESULT 6	
Q7XZ14	ID Q7XZ14 PRELIMINARY; PRT; 353 AA.
AC	Q7XZ14;
DT	01-OCT-2003 (TrEMBLrel. 25, Created)
DT	01-OCT-2003 (TrEMBurel. 25, Last sequence update)
DE	05-JUL-2004 (TrEMBurel. 27, Last annotation update)
DE	Transcription factor Dpl.
CN	Name=DPl;
OS	Populus tremula x Populus tremuloides.
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OX	eucosuridae I; Malpighiales; Salicaceae; Populus.
OX	NCBI_TaxID=47664;
[1]	
RP	SEQUENCE FROM N.A.
RA	Espinosa-Ruiz A.; Saxena S.; Schmidt J.; Mellerowicz E.; Bako L.S.;
RA	Bhalerao R.P.;
RA	Bhalerao R.P.;
RL	Submitted (JUN-2003) to the EMBL/GenBank/DDBJ databases.
[2]	
RP	SEQUENCE FROM N.A.
RA	Espinosa-Ruiz A.; Bhalerao R.P.;
RA	Submitted (MAY-2003) to the EMBL/GenBank/DDBJ databases.
[3]	
RP	SEQUENCE FROM N.A.
RA	Bhalerao R.P.; Sandberg G.;
RA	Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
CC	- - SUBCELLULAR LOCATION: Nuclear (By similarity).
CC	- - SIMILARITY: Belongs to the E2F/DP family.
DR	EMBL; AY307373; AAP73785.1; -.

```

DR EMBL; AF181998; AAQ13675.1; -.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 353 AA; 38444 MW; 33CF9F440D9D1F CRC64;

Query Match 36.1%; Score 476.5; DB 2; Length 353;
Best Local Similarity 42.7%; Pred. No. 5.1e-23;
Matches 112; Conservative 43; Mismatches 76; Indels 31; Gaps 6;

QY 2 APPRGAAATAALDLTGCHILEASSVPPPLPEAGNAVORVGVD-----GVHILEASSVPPPLPEAGNAVORVGK--AVDPDKR 52
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
46 SPSRSEHAWATPASDSTFURLNHLDIHADDAAT-----QAAANKKKKGQRAVGADK-- 99
QY 53 KKEKAAAPRITGWLREYSKIVCEKVEAKGRITTYNEVADEIYSELKSMHI-----GQGF 107
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
100 -----SGRLQFSIKVCEKVESKTTTYNEVADELVAEFADPSNSVSTPDQOY 149
QY 108 DEKNIRRVYDAFNVLIALRVIAKEKEIRWGLSNRYREKIKKLEVRKELVKNRK 167
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
150 DEKNIRRVYDALNVLNLMALDIISKDKKEIQWKGLPR-----TSMDSVEELK 209
QY 168 ALLOEIEKQFDLQNIKRLNOTLESSAENVN-GIRLPFLVKTSRKARVEIISDDSKFA 226
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
210 AYLOELEEQFVGLQNIQRNEQIYSSGNAPSGVSPFVLVQTRPHATVEVEISEDMQLV 269
QY 227 HPEFNGAPFTLHDDLSILEGVR 248
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
270 HFDNSTPFELHDDNYVVKAMK 291

RESULT 7
Q84VD5 ID Q84VD5 PRELIMINARY; PRT; 346 AA.
AC Q84VD5;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE DP TF.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
[1]
SEQUENCE FROM N.A.
MEDLINE=23111120; PubMed=14750518;
DOI=10.1023/B:PLAN.0000007001.30865.0f;
Cooper B., Hutchison D., Park S., Guimil S., Luginbuhl P., Ellero C.,
Goff S.A., Glazebrook J.;
"Identification of rice (Oryza sativa) proteins linked to the cyclin-
mediated regulation of the cell cycle.";
Plant Mol. Biol. 53:273-279(2003).
CC -!- SUBCELLULAR LOCATION: Nuclear (BY similarity).
CC -!- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AY224551; AA072671.1; -.
DR HSSP; Q14188; 1CF7.
DR Gramene; Q84VD5; -.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.

SQ SEQUENCE 379 AA; 40893 MW; 913F06C2975F80CB CRC64;

Query Match 35.4%; Score 466.5; DB 2; Length 379;
Best Local Similarity 37.8%; Pred. No. 2.5e-22;
Matches 111; Conservative 43; Mismatches 77; Indels 63; Gaps 8;

QY 3 PPRGGAAATAALDLTGCHILEASSVPP-----LPEAGNAVORVGKGVDP----- 48
Db :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
37 PPSGGAQASGSG-----GSAGSPSSRSEHQHVPAAAGNAAGAAAATPISNTFL 86

```



```

QY 49 -----DKDKKEKAAAPRITGWLREYSKIVCEKVEAKGRITTYNE 88
Db 87 RLNDLDIHGDAPSSQAPTSKXKRGARAVDPKGGRLGRQFMSKVKESKGRITTYNE 146
QY 89 VADEIYSELKSMARI-----GGFDEKIRRRVYDAFNVLIALRVIAKEKEIRW 138
Db 147 VADELVAEFADPNNSILPPDPDPNPAQQYDEKIRRRVYDALNVLAMAEIISKKEIQW 206
QY 139 MGLSNRYEKIKKLEVRKELV---NKIRNKALLQIEKQFDDLOQIKLRNOTLESSAE 195
Db 207 KGLPR---TSINDIEDLOTGLVGLKSRIEKNYTLQELQDFQVGMQKLIQNRNQLVSG- 262
QY 196 NV--NGIRLPFLVKTSSRKARVEISDDSKFAHFNFGAPFTLHDDLSILEGV 247
Db 263 NIPSGGVALPFILVQTRPHATVEISEDMQLVHFDNFSTPFELHDDSFVLKAM 316

RESULT 9
Q9LZE7 PRELIMINARY; PRT; 413 AA.
AC Q9LZE7;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Transcription factor-like protein.
GN NamesF12B4.160;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Bevan M., Pohl T., Weizenegger T., Bancroft I., Mewes H.W., Rudd S.,
RA Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -|- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AL162751; CAB83299.1; -.
DR PIR; T48364; T48364.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR009058; E2F_TDP.
DR InterPro; IPR003316; E2F_TDP.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 413 AA; 45604 MW; 239FDC6DAED2F723 CRC64;

Query Match 34.6%; Score 456; DB 2; Length 413;
Best Local Similarity 41.3%; Pred. No. 1.3e-21;
Matches 103; Conservative 41; Mismatches 68; Indels 34; Gaps 4;

QY 37 GNAVQKGAVDPPDKRKEKAAAPRITGWLREYSK-----IVCEKVE 79
Db 74 GDDAGSGAGSVKXKRGQRAAGPDKTGRGLRQFMSKGLISFAPIMLSKCLISCEKVE 133
QY 80 AKGRITTYNEVADEIYSELKSMARI-----QGFEKIRRRVYDAFNVLIALRVIAKEKK 134
Db 134 SKGRITTYNEVADELVAEFALPNNDGTSPOQQYDEKIRRRVYDALNVLAMAEIISKKEIKK 193
QY 135 EIRWGLSNRYEKIKKLEVRKELVKNKIKKALLQIEKQFDD-----LQNI 183
Db 194 EIOWGLPRTSLSDIBELKNRSLRNRIEKKTAYSQELSEQVMNIDITGLSASCLQNL 253
QY 184 KLRNQTLESSAENVN-GIRLPFLVKTSSRKARVEISDDSKFAHFNFGAPFTLHDDLS 242

```

```

Db 254 IQRNEHLYSSGNAPSGVALPFILVQTRPHATVEISEDMQLVHFDNFSTPFELHDDNF 313
QY 243 ILEGVR 248
Db 314 VLTKMK 319

RESULT 10
Q9LZ55 PRELIMINARY; PRT; 288 AA.
AC Q9LZ55;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE DP-2 transcription factor-like.
GN NamesF22P11.60;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Bevan M., Hilbert H., Braun M., Holzer E., Brandt A., Duesterhoeft A.,
RA Bancroft I., Mewes H.W., Rudd S., Lemcke K., Mayer K.F.X.;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
CC -|- SUBCELLULAR LOCATION: Nuclear (by similarity).
CC -|- SIMILARITY: Belongs to the E2F/DP family.
DR EMBL; AL162971; CAB85984.1; -.
DR PIR; T48268; T48268.
DR HSSP; Q14188; 1CF7.
DR GO; GO:0005634; C:nucleus; IEA.
DR GO; GO:0005667; C:transcription factor complex; IEA.
DR GO; GO:0003700; F:transcription factor activity; IEA.
DR GO; GO:0000074; P:regulation of cell cycle; IEA.
DR GO; GO:0006355; P:regulation of transcription, DNA-dependent; IEA.
DR InterPro; IPR003316; E2F_TDP.
DR InterPro; IPR009058; Wing_hlx_DNA_bnd.
DR Pfam; PF02319; E2F_TDP; 1.
DR DNA-binding; Nuclear protein; Transcription; Transcription regulation.
KW DNA-binding; Nuclear protein; Transcription; Transcription regulation.
SQ SEQUENCE 288 AA; 32560 MW; EC0AD7DC068F92B7 CRC64;

Query Match 32.5%; Score 428.5; DB 2; Length 288;
Best Local Similarity 41.3%; Pred. No. 5.2e-20;
Matches 88; Conservative 49; Mismatches 69; Indels 7; Gaps 2;

QY 40 VORKGAVDPPDKRKEKAAAPRITGWLREYSKIVCEKVEAKGRITTYNEVADEIYSELKS 99
Db 23 VRKLIIVDDSDSEIGSEKGGQSRSTGGGLRQFMSVQCQKLEAKKITTYKEVADEIISDFAT 82
QY 100 MAHIQ-----GFDEKIRRRVYDAFNVLIALRVIAKEKKEIRWGLSNRYEKIKKLE 153
Db 83 IQKNAEKPLNENEYKIRRRVYDALNVLAMAEIISDFATKKEIRWGLPITCKKQVEEVK 142
QY 154 EYRKELVKNIRNKALLQIEKQFDDLOQIKLRNOTLESSAEN-VNGIRLPFLVKTSSRK 212
Db 143 MDRNKVMSVQKKAAPFLKELREKVSLSMSRQEMVYKTCQGPASGFTLPFILLETPH 202
QY 213 ARVEIISDDSKFAHFNFGAPFTLHDDLSILE 245
Db 203 AVVEIISDDSKFAHFNFGAPFTLHDDLSILE 245

RESULT 11
Q9FNY3 PRELIMINARY; PRT; 292 AA.
AC Q9FNY3;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)

```

Query Match	32.5%	Score 428.5	DB 2	Length 292
Best Local Similarity	41.3%	Pred. No. 5.3e-20		
Matches 88	Conservative 49	Mismatches 69	Indels 7	Gaps 2
Qy	40	VORKGAVDPDKRRKEKAAAPRTGTWGLREYSKIYCEVKAEGTTTYNEVADEIYSELKS	99	
Db	27	VRRKKLIVDDDSIEIGSEKGQSRTSGGGIQRFQPSVMVCQKLEAKTKTYKEVADEIIISDFAT	86	
or	100	MAHTGO-----GFDEKNIRRVVDFAFNVIARLVIAKEKEIRWMGLSNRYVKIKKLE	153	

CC Comment=Additional isoforms seem to exist. Experimental
 CC confirmation may be lacking for some isoforms;
 CC Name=Alpha; Synonyms=49 kDa;
 CC IsoId=Q14188-1; Sequence=Displayed;
 CC Name=Beta; Synonyms=43 kDa;
 CC IsoId=Q14188-2; Sequence=VSP_001352, VSP_001353;
 CC Name=Gamma;
 CC IsoId=Q14188-3; Sequence=VSP_001352, VSP_001353, VSP_001354;
 CC Name=Delta; Synonyms=48 kDa;
 CC IsoId=Q14188-4; Sequence=VSP_001352;
 CC TISSUE SPECIFICITY: High levels in heart and skeletal muscle. Also
 CC found in placenta, kidney, brain, lung and liver. The presence as
 CC well as the abundance of the different transcripts appear to vary
 CC significantly in different tissues and cell lines.
 CC -!- PTM: Phosphorylated.
 CC -!- SIMILARITY: Belongs to the E2F/DP family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>)
 CC or send an email to license@isb-sib.ch.
 CC -----
 CC EMBL; L40386; AAA69016.1; -;
 CC EMBL; U18422; AAB60378.1; -;
 CC EMBL; BC021113; AAH21113.1; -;
 CC EMBL; U75488; AAB37321.1; -;
 CC EMBL; U35117; AAC50642.1; -;
 CC PDB; 1CF7; X-ray; B=121-215.
 CC TRANSFAC; T03000; -;
 CC H-InvDB; HIX0003733; -;
 CC Reactome; Q14188; -;
 CC MIM; 602160; -;
 CC GO; GO:0003677; F:DNA binding; TAS.
 CC GO; GO:0003712; P:transcription cofactor activity; TAS.
 CC InterPro; IPR003316; E2F_TDP.
 CC InterPro; IPR009058; Wing_hlx_DNA_bnd.
 CC Pfam; PF02319; E2F_TDP; 1.
 CC 3D-structure; Activator; Alternative splicing; Cell cycle;
 CC DNA-binding; Nuclear protein; Phosphorylation; Polymorphism;
 CC Transcription regulation.
 CC DOMAIN 103 118 Nuclear localization signal (By
 CC similarity).
 CC FT FT 129 210
 CC FT FT 176 210
 CC FT FT 219 292
 CC FT FT 229 261
 CC FT FT 274 330
 CC FT FT 432 446
 CC FT MOD RES 24 24
 CC FT MOD RES 42 42
 CC FT VARSPLIC 1 61
 CC FT FT
 CC FT VARSPLIC 103 118
 CC FT FT
 CC FT FT
 CC FT VARSPLIC 173 173
 CC FT FT
 CC FT VARIANT 64 64
 CC FT FT
 CC FT HELIX 131 146
 CC FT STRAND 148 149
 CC FT HELIX 151 163
 CC FT TURN 164 164
 CC FT TURN 166 167
 CC FT HELIX 170 172
 CC FT HELIX 174 193
 CC FT TURN 194 195
 CC FT TURN 197 197

FT STRAND 204 206
 SQ SEQUENCE 446 AA; 49236 MW; 19A6C85BAD61DFF1 CRC64;
 Query Match 29.9%; Score 394; DB 1; Length 446;
 Best Local Similarity 37.5%; Pred. No. 1.5e-17;
 Matches 94; Conservative 52; Mismatches 89; Indels 16; Gaps 8;
 QY 18 LTGVHILASSVPPLPEAGGNVQRKGAVDPPDKORKEKKAAPRITGWLREYSKIVCEK 77
 DB 89 VTQTHIAEATGWPP-----GDKRKARKFIDSFSESKSKKGDK-NGKLRHFSMKVCEK 142
 QY 78 VEAKGRRTVNEVADEIYSEL-KSMAHIG--QGFDENIRRRVYDAFNVLIALRVIAKEKK 134
 DB 143 VORKTTSYNEVADELVEFTNNHLLAADSAYDQKNIRRRVYDALNVLMMNNIISKKK 202
 QY 135 EIRWGLSNRYREYKTKLEVEKRVKLVKIRKQALLQELQKDFDLQNLKRNQTLSSA 194
 DB 203 EIKWIGLPTNSAQECQNLIEKQRIERIKQRAQLQELLQOIAFKNLVQRNRQNEQN 262
 QY 195 ENV---NGIRLPFVLVTSRKARVEIEISDDSKFAH-FEPNGAPFTLHDDLSILEGVRR 249
 DB 263 QGPPALNSTQIPFIIINTSRKTVIDCSISSD-KPEYLFNFENT-FEIHDDIEVLKRMGM 320
 QY 250 NSIGRAGRATL 260
 DB 321 SFGLESGRKSL 331
 RESULT 15
 Q8BHD2 PRELIMINARY; PRT; 385 AA.
 ID Q8BHD2;
 AC Q8BHD2;
 DT 01-MAR-2003 (TREMELrel. 23, Created)
 DT 01-MAR-2003 (TREMELrel. 23, Last sequence update)
 DT 25-OCT-2004 (TREMELrel. 28, Last annotation update)
 DE Mus musculus 10 days neonate skin cDNA, RIKEN full-length enriched
 DE library, clone:4732486C09 product:transcription factor Dp 2, full
 DE insert sequence (Mus musculus 16 days embryo lung cDNA, RIKEN full-
 DE length enriched library, clone:8430403A04 product:transcription factor
 DE Dp 2, full insert sequence).
 GN Name=A330080J22Rik;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung, and Skin;
 RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
 RA Carninci P., Hayashizaki Y.;
 RT "High-efficiency full-length cDNA cloning";
 RL Meth. Enzymol. 303:19-44(1999).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung, and Skin;
 RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
 RA RIKEN FANTOM Consortium;
 RT "Functional annotation of a full-length mouse cDNA collection";
 RL Nature 409:685-690(2001).
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung, and Skin;
 RA The FANTOM Consortium,
 RT the RIKEN Genome Exploration Research Group Phase I & II Team;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs";
 RL Nature 420:563-573(2002).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=C57BL/6J; TISSUE=Lung, and Skin;
 RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
 RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
 RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;

This Page Blank (uspto)

Db 28 VTQTHIAEAA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65
QY 78 VEAKGRTTYNEVADEIYSELKS-----MAHIGQGFEKNIIRRRVYDAFNVLIALRVIAKEK 133
Db 66 VQRKGTTSYNEVADELVSFTSNHNLAAADSQAYDQENIRRRVYDALNVLAMNIISEK 125
QY 134 KEIRWGLSNRYEYKIKLEEVKELVNKRKALLQIEKQFDLQNIKLNRQTLESS 193
Db 126 KEIKWIGLPTNSAQEQNLEIEKQRIERIKQKRAQLQELLQOIAFNKLVQRNRQEQ 185
QY 194 AEN---VNG-IRLPFVLVKTSRKARVEIEISDDSKFAH-FEFGAPFTLHDDLILSEGV 248
Db 186 NQPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNDT-FEIHDDIEVLKRMG 243
QY 249 RNSIGRAGRATL 260
Db 244 MSFGLSGKCSL 255

RESULT 2
US-09-189-627A-6
; Sequence 6, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; PRIOR FILING DATE: 1998-11-10
; PRIOR FILING DATE: 1996-09-30
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 370
; TYPE: PRT
; ORGANISM: mouse
US-09-189-627A-6

Query Match 29.5%; Score 389.5; DB 3; Length 370;
Best Local Similarity 37.3%; Pred. No. 4.7e-29;
Matches 94; Conservative 51; Mismatches 74; Indels 33; Gaps 8;
QY 18 LTGVHILEASSVPPLPEAGGNAVQRKGAVDPDKRKKKAAAPRITGWLGRYSKIVCEK 77
Db 28 VTQTHIAEAA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65
QY 78 VEAKGRTTYNEVADEIYSELKS-----MAHIGQGFEKNIIRRRVYDAFNVLIALRVIAKEK 133
Db 66 VQRKGTTSYNEVADELVSFTSNHNLAAADSQAYDQENIRRRVYDALNVLAMNIISEK 125
QY 134 KEIRWGLSNRYEYKIKLEEVKELVNKRKALLQIEKQFDLQNIKLNRQTLESS 193
Db 126 KEIKWIGLPTNSAQEQNLEIEKQRIERIKQKRAQLQELLQOIAFNKLVQRNRQEQ 185
QY 194 AEN---VNG-IRLPFVLVKTSRKARVEIEISDDSKFAH-FEFGAPFTLHDDLILSEGV 248
Db 186 NQPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNDT-FEIHDDIEVLKRMG 243
QY 249 RNSIGRAGRATL 260
Db 244 MSFGLSGKCSL 255

RESULT 3
US-09-710-861-6
; Sequence 6, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas

; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 6
; LENGTH: 370
; TYPE: PRT
; ORGANISM: mouse
US-09-710-861-6

Query Match 29.5%; Score 389.5; DB 3; Length 370;
Best Local Similarity 37.3%; Pred. No. 4.7e-29;
Matches 94; Conservative 51; Mismatches 74; Indels 33; Gaps 8;
QY 18 LTGVHILEASSVPPLPEAGGNAVQRKGAVDPDKRKKKAAAPRITGWLGRYSKIVCEK 77
Db 28 VTQTHIAEAA-----GWVPSKRSKKGDK-----NGKGLRHFSMKVCEK 65
QY 78 VEAKGRTTYNEVADEIYSELKS-----MAHIGQGFEKNIIRRRVYDAFNVLIALRVIAKEK 133
Db 66 VQRKGTTSYNEVADELVSFTSNHNLAAADSQAYDQENIRRRVYDALNVLAMNIISEK 125
QY 134 KEIRWGLSNRYEYKIKLEEVKELVNKRKALLQIEKQFDLQNIKLNRQTLESS 193
Db 126 KEIKWIGLPTNSAQEQNLEIEKQRIERIKQKRAQLQELLQOIAFNKLVQRNRQEQ 185
QY 194 AEN---VNG-IRLPFVLVKTSRKARVEIEISDDSKFAH-FEFGAPFTLHDDLILSEGV 248
Db 186 NQPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNDT-FEIHDDIEVLKRMG 243
QY 249 RNSIGRAGRATL 260
Db 244 MSFGLSGKCSL 255

RESULT 4
US-08-723-415B-8
; Sequence 8, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: LaThangue, Nicholas B.
; APPLICANT: deLaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996


```
/ TITLE OF INVENTION: THEREOF
/ NUMBER OF SEQUENCES: 21
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: NIXON & VANDERHYE P.C.
/ STREET: 1100 No. 5859199th Glebe Rd. 8th floor
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22201-4741
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: PatentIn Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/723,415B
/ FILING DATE: 30-SEP-1996
/ CLASSIFICATION: 435
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: GB 9610195.1
/ FILING DATE: 15-MAY-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: Crawford, Arthur R.
/ REGISTRATION NUMBER: 25,327
/ REFERENCE/DOCKET NUMBER: 117-220
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-816-4000
/ TELEFAX: 703-816-4100
/ INFORMATION FOR SEQ ID NO: 2:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 446 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-723-415B-2

Query Match 29.5%; Score 389; DB 2; Length 446;
Best Local Similarity 38.6%; Pred. No. 6.9e-29;
Matches 98; Conservative 48; Mismatches 86; Indels 22; Gaps 10;

QY 18 LTGVHILEASSVPPLPEAGNAVQRKGA---VDPDKDRKKEKAAAPRITGWLREYSKIV 74
Db 89 VTQTHIAEA-----AGWVPSDRKAREFIDSFESKRSKKGDK-NGKGLRHFSMKV 139

QY 75 CEKVEAKGRTTYNEVADEIYSEL-KSMAHIG--QGFDEKNIRRVYDAFNVLIALRVIK 131
Db 140 CEKVEAKGRTTYNEVADEIYSEL-KSMAHIG--QGFDEKNIRRVYDAFNVLIALRVIK 131
QY 132 EKKEIRMMGLSNRYEYKIKLEEVKELVNKIRNKALLQEIKEKQFDDLQNIKLNRQTL 191
Db 200 EKKEIRMMGLSNRYEYKIKLEEVKELVNKIRNKALLQEIKEKQFDDLQNIKLNRQTL 191
QY 192 SSAEN---VNG-IRLPFVLVKTARKARVEIEISDDSKFAH--FEFGAPFTLHDDLIL 246
Db 260 QNQGGPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFONT-FEIHDDIEVLKR 317

QY 247 VRNSIGRAGRATL 260
Db 318 MGMSFGLSGKCSL 331

RESULT 9
US-09-710-861-2
/ Sequence 2, Application US/09710861
/ Patent No. 6387649
/ GENERAL INFORMATION:
/ APPLICANT: La Thangue, Nicholas
/ APPLICANT: de la Luna, Susana
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
/ FILE REFERENCE: 620-54
/ CURRENT APPLICATION NUMBER: US/09/710,861
/ PRIOR FILING DATE: 2000-11-13
/ PRIOR APPLICATION NUMBER: US/09/189,627
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR APPLICATION NUMBER: 08/723,415
/ PRIOR FILING DATE: 1996-09-30
/ PRIOR APPLICATION NUMBER: GB 9610195
/ PRIOR FILING DATE: 1996-05-15
/ NUMBER OF SEQ ID NOS: 25
/ SOFTWARE: PatentIn Ver. 2.0
/ SEQ ID NO 2
/ LENGTH: 446
/ TYPE: PRT
/ ORGANISM: mouse
/ US-09-710-861-2

Query Match 29.5%; Score 389; DB 3; Length 446;
Best Local Similarity 38.6%; Pred. No. 6.9e-29;
Matches 98; Conservative 48; Mismatches 86; Indels 22; Gaps 10;

QY 18 LTGVHILEASSVPPLPEAGNAVQRKGA---VDPDKDRKKEKAAAPRITGWLREYSKIV 74
Db 89 VTQTHIAEA-----AGWVPSDRKAREFIDSFESKRSKKGDK-NGKGLRHFSMKV 139

QY 75 CEKVEAKGRTTYNEVADEIYSEL-KSMAHIG--QGFDEKNIRRVYDAFNVLIALRVIK 131
Db 140 CEKVEAKGRTTYNEVADEIYSEL-KSMAHIG--QGFDEKNIRRVYDAFNVLIALRVIK 131
QY 132 EKKEIRMMGLSNRYEYKIKLEEVKELVNKIRNKALLQEIKEKQFDDLQNIKLNRQTL 191
Db 200 EKKEIRMMGLSNRYEYKIKLEEVKELVNKIRNKALLQEIKEKQFDDLQNIKLNRQTL 191
QY 192 SSAEN---VNG-IRLPFVLVKTARKARVEIEISDDSKFAH--FEFGAPFTLHDDLIL 246
Db 260 QNQGGPPAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFONT-FEIHDDIEVLKR 317

QY 247 VRNSIGRAGRATL 260
Db 318 MGMSFGLSGKCSL 331

RESULT 8
US-09-189-627A-2
/ Sequence 2, Application US/09189627A
/ Patent No. 6159691
/ GENERAL INFORMATION:
/ APPLICANT: La Thangue, Nicholas
/ APPLICANT: de la Luna, Susana
/ TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
/ FILE REFERENCE: 620-54
/ CURRENT APPLICATION NUMBER: US/09/189,627A
/ PRIOR FILING DATE: 1998-11-10
/ PRIOR APPLICATION NUMBER: 08/723,415
```

QY 132 EKKEIRWGLSNRYEYKIKLEVRKLVNKNRKKALLQEIIEKQFDDLNQIKLRNQTLE 191
Db 200 EKKEIKWGLPTNSAQEONLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNQRNE 259
QY 192 SSAEN---VNG-IRLPLVLTSTRKARVEIEISDDSKFAH-FEFGAPFTLHDDLSILEG 246
Db 260 QOQGPVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNT-FEIHDDIEVLKR 317
QY 247 VRNSIGRATL 260
Db 318 MGMSFGLESKCSL 331

RESULT 10
US-09-949-016-9220
; Sequence 9220, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9220
; LENGTH: 331
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9220

Query Match 29.2%; Score 385; DB 4; Length 331;
Best Local Similarity 38.7%; Pred. No. 1.1e-28;
Matches 92; Conservative 48; Mismatches 84; Indels 14; Gaps 8;

QY 34 EAGNAVQRGA---VDPDKRKKKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVA 90
Db 67 ESQSNKSDRKARKFIDSFSESKSKGDK-NGKGLRHFMSKVCVKQKRTTTSNEVA 125
QY 91 DEIYSEL-KSMAHIG--QGDEKNIRRVYDAFNVLIALRVIAKEKEIRWGLSNRYE 147
Db 126 DELVSEFTSNHAAADSAYDQKNIRRVYDALNVLMMNIIISKEKEIKWGLPTNSAQ 185
QY 148 KIKLEVRKLVNKNRKKALLQEIIEKQFDDLNQIKLRNQTLESSAENV---NGIRLP 203
Db 186 EONLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNQRNEQOQGPVAVNSTIQLP 245
QY 204 FVLVLTSTRKARVEIEISDDSKFAH-FEFGAPFTLHDDLSILEGVRNSIGRATL 260
Db 246 FIIINTSRKTVIDCSISSD-KFEYLFNFNT-FEIHDDIEVLKRMGMSFGLESKCSL 301

RESULT 11
US-08-723-415B-4
; Sequence 4, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: Lathangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Giebe Rd. 8th floor
; CITY: Arlington

STATE: VA
COUNTRY: USA
ZIP: 22201-4741
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: IBM PC compatible
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/723,415B
FILING DATE: 30-SEP-1996
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 9610195.1
FILING DATE: 15-MAY-1996
ATTORNEY/AGENT INFORMATION:
NAME: Crawford, Arthur R.
REGISTRATION NUMBER: 25,327
REFERENCE/DOCKET NUMBER: 117-220
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-816-4000
TELEFAX: 703-816-4100
INFORMATION FOR SEQ ID NO: 4:
SEQUENCE CHARACTERISTICS:
LENGTH: 369 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-723-415B-4

Query Match 29.1%; Score 384; DB 2; Length 369;
Best Local Similarity 37.5%; Pred. No. 1.6e-28;
Matches 94; Conservative 50; Mismatches 75; Indels 32; Gaps 9;

QY 18 LTCVHILEASSVPPPLPEAGNAVQRKGVDPDKRKKKAAAPRITGWLREYSKIVCEK 77
Db 28 VTQTHIAEAA-----GWVPSKRSKKGDK-----NGKGLRHFMSKVCVK 65
QY 78 VEAKGRTTYNEVADEIYSEL-KSMAHIG--QGDEKNIRRVYDAFNVLIALRVIAKEK 134
Db 66 VOKGRTTYSNEVADELVSFTSNHAAADSAYDQKNIRRVYDALNVLMMNIIISKEK 125
QY 135 EIRWGLSNRYEYKIKLEVRKLVNKNRKKALLQEIIEKQFDDLNQIKLRNQTLESSA 194
Db 126 EIKWGLPTNSAQEONLEIKQRIERIKQKRAQLQELLQOIAFKNLVQRNQRNEQON 185
QY 195 EN---VNG-IRLPLVLTSTRKARVEIEISDDSKFAH-FEFGAPFTLHDDLSILEGVR 249
Db 186 QGPPAVAVNSTIQLPFIINTSRKTVIDCSISSD-KFEYLFNFNT-FEIHDDIEVLKRMGM 243
QY 250 NSIGRATL 260
Db 244 SFGLESKCSL 254

RESULT 12
US-09-189-627A-4
; Sequence 4, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; CURRENT FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4

```

Db      186 QCPAVNSTIQPFIIINTSRKTVDCSISD-KFYLNFONT-FEIHDDIEVLKMGCM 243

QY      250 NSTGRAGRATL 260
          :|: |: |
Db      244 SFLESGKCSSL 254

RESULT 14
US-08-723-415B-11
; Sequence 11, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: Lathague, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHVE P.C.
; STREET: 1100 NO. 5859199TH Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-723-415B-11

Query Match      28.7%; Score 378.5; DB 2; Length 410;
Best Local Similarity 42.4%; Pred. No. 6.2e-28;
Matches      87; Conservative 41; Mismatches 60; Indels 17; Gaps 7

QY      50 KDRKEAAPRIITGWLREYSKIVCEKVAKRGRTTNEVADEIYSEUKSM-AHI---GQ 105
          :||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||
Db      105 ENRXGEK-----NGKLRFHSKMVCKEVQRGTTSYNEVADLVAFPSAADNHILPNES 158
          :||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||

QY      106 GFDEKNTRRRYYDAFNULIALRVIAKEKETRWGLSNRYREKIKKLEVRKELVNKIRN 165
          :||| ||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||
Db      159 AYDKNTRRRRYYDALNVLMANNIIISKETIKWIGLPFTNSAQEQCNLEVERQRRLERIKQ 218
          :||| ||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||

QY      166 KKALLOETEKQFDLLONIKLRNOTLESSAENV----NGIRLPFFVLVKTSRKARVEIISD 221
          :||| ||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||
Db      219 KQSQLQEILIQIAFKNLVQRNQAEQAQQRPSPNSVIHLFPFIINVTSRKTVIDCSISN 278

QY      222 DSKFAH-FEFNGAPPTLHDLSILE 245
          ||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||
Db      279 D-KFEYLFNFONT-FEIHDDIEVLK 301
          ||| ||| :||| ||| :||| ||| :||| ||| :||| ||| :|||

```

RESULT 15
US-08-428-131-2
; Sequence 2, Application US/08428131
; Patent No. 5863757
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas Barrie
; TITLE OF INVENTION: Transcription Factor DP-1
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nixon & Vanderhye
; STREET: 1100 No. 5863757th Glebe Road, 8th Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22201-4714
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25 (BPO)
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/428,131
; FILING DATE: 23-JUN-1995
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Arthur R. Crawford
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-181
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 816-4000
; TELEFAX: (703) 816-4100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 410 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-428-131-2

Query Match 28.7%; Score 378.5; DB 2; Length 410;
Best Local Similarity 42.4%; Pred. No. 6.2e-28;
Matches 87; Conservative 41; Mismatches 60; Indels 17; Gaps 7;

QY 50 KKKKKKAAAPRTGTGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSM-AHI---GQ 105
Db 105 RRRKGEK-----NGKGLRHFSMKVCKVKQKRTTTSYNEVADELVAEFGAADNHILPNES 158

QY 106 GFDEKNIRRVYDAFNVLIARVIAKEKEIRMMGLSNRYEYKIKKLEEVYRKELVKNKRN 165
Db 159 AYDQKNIRRVYDALNVLMAMNIIKKEKEIKWIGLPTNSAQECQCNLEVERQRRLEIKQ 218

QY 166 KKALLOEIEKQFDDLQNIKLRNQTLESSAENV----NGRLPVLVKTSRKARVEIISD 221
Db 219 KOSQLOELIQQIAFKNLVQRNRAEQARRPPPPNSVIHLPIIVTSTRKTVDCISGN 278

QY 222 DSKFAH-FEFNGAPFTLHDDLSILE 245
Db 279 D-KFEYLFNFNT-FEIHDDIEVK 301

Search completed: February 4, 2005, 20:34:32
Job time : 44 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 4, 2005, 20:33:05 ; Search time 132 Seconds
(without alignments)
644.003 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGGAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1373511 seqs, 325702437 residues

Total number of hits satisfying chosen parameters: 1373511

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA:*

1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/prodata/2/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/prodata/2/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/prodata/2/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/prodata/2/pubpaa/US10D_PUBCOMB.pep:*
17: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep:*
18: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep:*
19: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep:*
20: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	842	63.8	369	16 US-10-437-963-136371	Sequence 136371,
2	464	35.2	336	15 US-10-425-114-46555	Sequence 46555, A
3	453.5	34.4	320	15 US-10-424-599-186648	Sequence 186648,
4	438	33.2	314	15 US-10-424-599-185947	Sequence 185947,
5	374.5	28.4	355	14 US-10-106-698-4846	Sequence 4846, A
6	374	28.4	120	16 US-10-767-701-49206	Sequence 49206, A
7	372.5	28.2	318	15 US-10-437-963-166158	Sequence 166158,
8	361	27.4	222	15 US-10-425-114-36974	Sequence 36974, A
9	354	26.8	575	9 US-09-220-091-7	Sequence 7, Appli
10	342.5	26.0	207	15 US-10-425-114-71403	Sequence 71403, A
11	337.5	25.6	263	16 US-10-437-963-167076	Sequence 167076,
12	299.5	22.7	405	14 US-10-053-248-24	Sequence 24, Appl
13	299.5	22.7	405	16 US-10-345-837-24	Sequence 24, Appl

14	225.5	17.1	165	15	US-10-424-599-234773	Sequence 234773,
15	203.5	15.4	74	14	US-10-214-188-10	Sequence 10, Appl
16	150	11.4	134	16	US-10-437-963-166159	Sequence 166159
17	133	10.1	37	9	US-09-900-147-1	Sequence 1, Appli
18	113.5	8.6	250	15	US-10-282-122A-46532	Sequence 46532, A
19	112.5	8.5	1156	15	US-10-369-493-43	Sequence 43, Appl
20	111.5	8.5	1178	15	US-10-282-122A-52434	Sequence 52434, A
21	111	8.4	444	15	US-10-389-566-435	Sequence 435, App
22	110.5	8.4	123	15	US-10-424-599-257840	Sequence 257840,
23	110	8.3	270	15	US-10-424-599-248849	Sequence 248849,
24	108.5	8.2	1957	15	US-10-369-493-2070	Sequence 2070, Ap
25	107	8.1	718	16	US-10-437-963-196494	Sequence 196494,
26	106	8.0	397	15	US-10-389-566-691	Sequence 691, App
27	105	8.0	317	15	US-10-389-566-692	Sequence 692, App
28	104.5	7.9	681	16	US-10-767-701-43781	Sequence 43781, A
29	104.5	7.9	965	10	US-09-842-484A-2	Sequence 2, Appli
30	104.5	7.9	965	10	US-09-842-484A-4	Sequence 4, Appli
31	104.5	7.9	965	14	US-10-184-485-3	Sequence 3, Appli
32	104.5	7.9	965	14	US-10-217-613-3	Sequence 3, Appli
33	104.5	7.9	965	14	US-10-217-613-8	Sequence 8, Appli
34	104.5	7.9	965	16	US-10-642-248-4	Sequence 4, Appli
35	103.5	7.8	324	9	US-09-220-091-9	Sequence 9, Appli
36	103	7.8	450	15	US-10-369-493-13626	Sequence 13626, A
37	102.5	7.8	563	15	US-10-425-114-71965	Sequence 71965, A
38	102.5	7.8	1069	15	US-10-424-599-259331	Sequence 259331,
39	101.5	7.7	972	9	US-09-879-959-10	Sequence 10, Appl
40	101.5	7.7	972	14	US-10-011-768B-9	Sequence 9, Appli
41	101.5	7.7	972	14	US-10-011-771B-9	Sequence 9, Appli
42	101.5	7.7	972	14	US-10-172-527-10	Sequence 10, Appl
43	101.5	7.7	972	14	US-10-326-185-95	Sequence 95, Appl
44	101.5	7.7	972	15	US-10-309-560-8	Sequence 8, Appli
45	101.5	7.7	972	16	US-10-642-248-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-10-437-963-136371
; Sequence 136371, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 136371
; LENGTH: 369
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_37957C.1.p.ep
US-10-437-963-136371

Query Match 63.8%; Score 842; DB 16; Length 369;
Best Local Similarity 50.4%; Pred. No. 2e-66;
Matches 185; Conservative 32; Mismatches 36; Indels 114; Gaps 5;

QY 1 MAPPRGGAATAALDLTG.....SILEGVRNRSIGRAGRATLH-----34

DB 1 MAPPGDAAAAAAGPLANLIREGAGLPSRPPRPTSDSFAPISREGDDIPQ 60

QY 35 -----AGGNAVQRK--GAVDPDKRKKKAAAPRITGWLREYSKI-----73

```
Db 61 KKSVALRSGGGNAAREEGGA---NRCKEKTGAQRITGWLRFSGKIGTKLPDPL 117
QY 74 -----VCEKVEAKGRTTYN 87
Db 118 SRPGKGHSYCAEVAVFTLTPYGGKPGNGRRAVTGGWVPGVDFVSKKVEAKGRTTYN 177
QY 88 E-----VADEIYSELKSMAGHIGGPFDEKIRRRVYDAFNVLIALRVI 129
Db 178 EIMVQTSNDVYTSSELIVADEIFAELKSITQNGLEFDEKIRRRVYDAFNVLIALRVI 237
QY 130 AKEKKEIRWGLSNRYEYKKELEVRKELVNKIRNKKALLOBIEKQFDLQNIKLNRQT 189
Db 238 AKDKKEIRWGLSNRYEYKKELEVRKELVNKIRNKKALLOBIEKQFDLQNIKLNRQA 297
QY 190 LESSAENVGIRLPVLVKTSKARVEIBISDDSKFAHFEFNGAPFTLHDDLILGVRR 249
Db 298 SQPAESVNGILLPFLINTSKARVEIBISEDSKFAHFEFNGAPFTMDDVSIIEAIRR 357
QY 250 NSIGRAG 256
Db 358 NNKGRAG 364
```

RESULT 2

```
US-10-425-114-46555
; Sequence 46555, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 46555
; LENGTH: 336
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: 700347688_FLI.pep
US-10-425-114-46555
```

```
Query Match 35.2%; Score 464; DB 15; Length 336;
Best Local Similarity 39.9%; Pred. No. 8e-33;
Matches 107; Conservative 45; Mismatches 76; Indels 40; Gaps 6;

QY 5 RGGAAAAATAALDLTGWHILEASSVPLPEAGGNVQKGAVDPKD-----RKX 54
Db 22 RGGPAPSASASV-----TPASESTVARRLNGLDPHGDDAPSSQPATSKKK 67
QY 55 E---KAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMAGH----- 103
Db 68 KRGARAVGPDKNRGLRQFSMRVCEKVESKGRRTTYNEVADELVAEF--SDPNINIDSPDPD 126
QY 104 ---GGGFDEKIRRRVYDAFNVLIALRVIAKEKEIRWGLSNRYEYKKELEVRKELV 160
Db 127 NSSAQYDEKIRRRVYDALNVLMAEIIISKQKEIQWRGLPKTSWNDIEELKTEVNGLK 186
QY 161 NKIRNKKALIQEIKQFDLQNIKLNRQTLESSAENVN--GIRLPVLVKTSKARVEIEI 219
Db 187 GRIDKKNAYLQDQYACIQLNVRNEQLYSGDAPSGGVALPFIIVQTRPHATVEVEI 246
QY 220 SDDSKFAHFEFNGAPFTLHDDLILGV 247
Db 247 SEDMQHVFDFNSTPTPELOQDSFVLKAM 274
```

RESULT 3

```
US-10-424-599-186648
; Sequence 186648, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 186648
; LENGTH: 320
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(320)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_139556C.1.pep
US-10-424-599-186648
```

```
Query Match 34.4%; Score 453.5; DB 15; Length 320;
Best Local Similarity 42.6%; Pred. No. 6.5e-32;
Matches 104; Conservative 42; Mismatches 73; Indels 25; Gaps 5;

QY 27 SSVPP-----LPEAGGNVQKGAVDPKDKRKEKAAAPRITGWLREYS 71
Db 21 NSLPPVPDKTILKLNHLVDHADDAGSHASLGS---KKKRGQRAVGDPDKSGRGLRQFS 77
QY 72 KIVCEKVEAKGRTTYNEVADEIYSELKSMAGH-----GGGFDEKIRRRVYDAFNVLIAL 126
Db 78 MKVCEKVESGRTTYNEVADELVAEFSEPSNSLPELDPQOQYDEKIRRRVYDALNVLMA 137
QY 127 RVIAKEKEIRWGLSNRYEYKKELEVRKELVNKIRNKKALLOBIE--KQFDDLQNIKL 185
Db 138 DIISKDKKEIQWRGLPRTTVDIEELKTERLGLNRRIEKKTAYLOELEXEQFTGLQNLIQ 197
QY 186 RNQTLSSAENVNG--IRLPVLVKTSKARVEIEISDDSKFAHFEFNGAPFTLHDDLIL 244
Db 198 RNEQLYSSRNPPSGSVSLPFIIVQTRPHATGGMSEISDMQLVHDFDNSTPFELHDDNYVL 257
QY 245 EGVV 248
Db 258 KAMK 261
```

RESULT 4

```
US-10-424-599-185947
; Sequence 185947, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 185947
; LENGTH: 314
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
```



```
; OTHER INFORMATION: Clone ID: PAT_MRT3847_138923C.1.pap
US-10-424-599-185947

Query Match      33.2%; Score 438; DB 15; Length 314;
Best Local Similarity 45.9%; Pred. No. 1.5e-30;
Matches 94; Conservative 39; Mismatches 66; Indels 6; Gaps 2;

QY 50 KDRKKEKAAAPRITGWLREYSKIVCEKAEKGRRTTYNEVADEIYSELKSMAH-----IG 104
DB 54 KRRGGQAGGPKDSGRCLAYFYMVKVCEKESRGRSYNEVADELVAEPADPINGSVTPDQ 113

QY 105 QGFDEKNIRRVYDAFNVLIARLVIAKEKEIRWGLSNRYRYEKIKKLEVRKELVNKIR 164
DB 114 QQYDEKNIRRVYDALNVLMAMDIIISKKEIQWGLPRTSLSDIBELKSERLGLNRIE 173

QY 165 NKALLOEIEKQFDDLOKIKLRNQTLESSAENVN-GIRLPFLVKTSRKARVEIISDD 223
DB 174 KKAAYLQELSEEQYIGLQKIQORNEQYSSGNAPNGGVSLFVLQTRPHATVEVEISEDM 233

QY 224 KFAHFEFGAPFTLHDDLSTILEGVR 248
DB 234 QLVHFDYICAPFEMHDDNVVLKAMK 258

RESULT 5
US-10-106-698-4846
; Sequence 4846, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent in Ver. 3.0
; SEQ ID NO 4846
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (342)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (348)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (351)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: MISC FEATURE
; LOCATION: (352)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-106-698-4846

Query Match      28.4%; Score 374.5; DB 14; Length 355;
Best Local Similarity 42.0%; Pred. No. 8.1e-25;
Matches 86; Conservative 42; Mismatches 60; Indels 17; Gaps 7;

QY 50 KDRKKEKAAAPRITGWLREYSKIVCEKAEKGRRTTYNEVADEIYSELKSM-AHI---GQ 105
DB 111 RNRKGEK-----NGKGLRHFSMKVCEKQKQRTTSYNEVADELVAEFSADNHLPNES 164

QY 106 GFDEKNIRRVYDAFNVLIARLVIAKEKEIRWGLSNRYRYEKIKKLEVRKELVNKIR 165
DB 165 AYDQKNIRRVYDALNVLMAMNIIISKEKEIKWIGLPTNSAQBCQNLEVRQRRLRIKQ 224
```

```
QY 166 KKALLOEIEKQFDDLOKIKLRNQTLESSAENV-----NGIRLPFLVKTSRKARVEISD 221
DB 225 KQSQLELILQQAIFAKNLVQRNRHAEQQASRPPPPNSVIHLPIIVNTSKTKTVIDCSISN 284

QY 222 DSKFAH-FEFGAPFTLHDDLSTILE 245
DB 285 D-KFEYLFNFNDNT-FEIHDDIEVLK 307

RESULT 6
US-10-767-701-49206
; Sequence 49206, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 49206
; LENGTH: 120
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-039-P1-K1-D2.pap
US-10-767-701-49206

Query Match      28.4%; Score 374; DB 16; Length 120;
Best Local Similarity 67.5%; Pred. No. 2e-25;
Matches 77; Conservative 17; Mismatches 18; Indels 2; Gaps 2;

QY 138 WMGLSNRYRYEKIKKLEVRKELVNKIRNKKALLQIEIKQFDDLOKIKLRNQTLESSAENV 197
DB 2 WACL-HFOYEYI-KLEETRKMIRVKNKKLQIEIERQFDDLOKIKLRNQTLESSAENV 59

QY 198 NGIRLPFLVKTSRKARVEIISDDSKFAHFEFGAPFTLHDDLSTILEGVRNS 251
DB 60 NGICLPFLLVKASRKARVEIISENSKFAGFDPNCTPFTLHDDVSILEAIRCNN 113

RESULT 7
US-10-437-963-166158
; Sequence 166158, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 166158
; LENGTH: 318
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_64895C.1.pap
US-10-437-963-166158

Query Match      28.2%; Score 372.5; DB 16; Length 318;
Best Local Similarity 37.7%; Pred. No. 1e-24;
```

RESULT 9
US-03-220-031-7
; Sequence 7, Application US/09220091
; Patent No. US20020064523A1
; GENERAL INFORMATION:
; APPLICANT: H. Robert Horvitz
; APPLICANT: Craig Cecil
; APPLICANT: Xiaowei Lu

	Query Match	26.0%;	Score 342.5;	DB 15;	Length 207;
	Best Local Similarity	44.4%;	Pred. No. 2.7e-22;		
	Matches	75;	Mismatches 48;	Indels 17;	Gaps 4;
Qy	105	QGFDEKNIRRVYDAFNVILALRVIAEKKEIRWGLSNRYEKIKLEEVKELV---	N 161		
		: : :			
Db	1	QQYDEKNIRRVYDALNVLMAMDIISKDKKEIQWGLPR---	TSISDIEEMKTELVLGLKG 57		
Qy	162	KIRNKKALLOBIEQOFDDLQNIKLNR--QTLESSAENVNGIRLPFLVKTSTKARVEIET	220		
		: : :			
Db	58	RIEKSAYLOBOQYVGLQNLIOQRNEQSYSGNTPSGVALPFLVOTRPHPTVEVEIS	117		
		: : :			

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-248-24

Query Match      22.7%; Score 299.5; DB 14; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCRLSMKMWETVQKGTTCQEVVGVGVAKFRAASNHASPNEA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAWNIISREKKIKWIGLTNTSAQNCONLVERQKRLERIKQK 214
QY 167 KALLQIEIKQFDDLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRQVVEQVQRPLPNSVIHVPFIIISSSKKTVCNCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 13
US-10-345-837-24
; Sequence 24, Application US/10345837
; Publication No. US20040137440A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Nucleic Acid
; FILE REFERENCE: P-IS 5589
; CURRENT APPLICATION NUMBER: US/10/345,837
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-345-837-24

Query Match      22.7%; Score 299.5; DB 16; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCRLSMKMWETVQKGTTCQEVVGVGVAKFRAASNHASPNEA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAWNIISREKKIKWIGLTNTSAQNCONLVERQKRLERIKQK 214
QY 167 KALLQIEIKQFDDLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRQVVEQVQRPLPNSVIHVPFIIISSSKKTVCNCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 14
US-10-424-599-234773
; Sequence 24, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-248-24

Query Match      22.7%; Score 299.5; DB 14; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCRLSMKMWETVQKGTTCQEVVGVGVAKFRAASNHASPNEA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAWNIISREKKIKWIGLTNTSAQNCONLVERQKRLERIKQK 214
QY 167 KALLQIEIKQFDDLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRQVVEQVQRPLPNSVIHVPFIIISSSKKTVCNCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 13
US-10-345-837-24
; Sequence 24, Application US/10345837
; Publication No. US20040137440A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Nucleic Acid
; FILE REFERENCE: P-IS 5589
; CURRENT APPLICATION NUMBER: US/10/345,837
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-345-837-24

Query Match      22.7%; Score 299.5; DB 16; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCRLSMKMWETVQKGTTCQEVVGVGVAKFRAASNHASPNEA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAWNIISREKKIKWIGLTNTSAQNCONLVERQKRLERIKQK 214
QY 167 KALLQIEIKQFDDLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRQVVEQVQRPLPNSVIHVPFIIISSSKKTVCNCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 14
US-10-424-599-234773
; Sequence 24, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

```
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-053-248-24

Query Match      22.7%; Score 299.5; DB 14; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCRLSMKMWETVQKGTTCQEVVGVGVAKFRAASNHASPNEA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAWNIISREKKIKWIGLTNTSAQNCONLVERQKRLERIKQK 214
QY 167 KALLQIEIKQFDDLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRQVVEQVQRPLPNSVIHVPFIIISSSKKTVCNCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 13
US-10-345-837-24
; Sequence 24, Application US/10345837
; Publication No. US20040137440A1
; GENERAL INFORMATION:
; APPLICANT: Lin, Biaoyang
; TITLE OF INVENTION: Androgen Regulated Nucleic Acid
; FILE REFERENCE: P-IS 5589
; CURRENT APPLICATION NUMBER: US/10/345,837
; PRIOR FILING DATE: 2003-01-15
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 24
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-345-837-24

Query Match      22.7%; Score 299.5; DB 16; Length 405;
Best Local Similarity 33.7%; Pred. No. 4.6e-18;
Matches 68; Conservative 47; Mismatches 72; Indels 15; Gaps 4;

QY 51 DRKKEKAAAPRITGWLREYSKIVCEKVEAKGRTTYNEVADEIYSELKSMHIG----QG 106
Db 101 NRKGEK-----NGMGLCRLSMKMWETVQKGTTCQEVVGVGVAKFRAASNHASPNEA 154
QY 107 FDEKNIRRRYDAFNVLIALRVIAKEKKEIRWMGLSNRYEKIKGLEEVKELVKNIRNK 166
Db 155 YDVKNIKRRTYDALNVLMAWNIISREKKIKWIGLTNTSAQNCONLVERQKRLERIKQK 214
QY 167 KALLQIEIKQFDDLQNIKLNRNOTLESSAEN-----VNGIRLPFVLVKTSRKARVEIEISDD 222
Db 215 QSELOQLILQIATFKNLVLNRQVVEQVQRPLPNSVIHVPFIIISSSKKTVCNCSISDD 274
QY 223 SKFAHFEENGAPFTLHDDLIL 244
Db 275 KSEYLPKFNSS-FEIHDDTEVL 295

RESULT 14
US-10-424-599-234773
; Sequence 24, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
```

APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 234773
LENGTH: 165
TYPE: PRT
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MRT3847_54029C.1.pep
US-10-424-599-234773

Query Match 17.1%; Score 225.5; DB 15; Length 165;
Best Local Similarity 39.9%; Pred. No. 5.2e-12; Indels 23; Gaps 3;
Matches 55; Conservative 20; Mismatches 40;

QY 1 MAPPRGMAAAATAALDLTGWHILEASSVPPLPEAGGNAVORKGAVDPDKDRKKKAAAP 60
DB 46 MATPADSTFLRLNHLIDHG-----DDAGSQGTVAK-----KKKGQRAVGS 87

QY 61 RITGWGLREYKIVCEKVEAKGRTTYNEVADEIYSELKSMAH-----IGQGFDEKNIRRR 115
DB 88 DKSGRGLRQFSMKVCEKVESKGRRTTYNEVADELVAEFADPDPINGVSTPDQOQYDEKNIRRR 147

QY 116 VYDAFNVLIALRVIAKEK 133
DB 148 VYDALNVLWAMDIIISKDK 165

RESULT 15
US-10-214-188-10
Sequence 10, Application US/10214188
Publication No. US20030022260A1
GENERAL INFORMATION:
APPLICANT: LA THANGUE, NICHOLAS B.
HJUNANS, ELEANORE M.
TITLE OF INVENTION: TRANSCRIPTION FACTOR E2F-5
NUMBER OF SEQUENCES: 25
CORRESPONDENCE ADDRESS:
ADDRESSEE: NIXON & VANDERHYE P.C.
STREET: 1100 NORTH GLEBE ROAD
CITY: ARLINGTON
STATE: VIRGINIA
COUNTRY: U.S.A.
ZIP: 22201-4714
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/214,188
FILING DATE: 08-Aug-2002
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US/08/894,139
FILING DATE: 13-AUG-1997
ATTORNEY/AGENT INFORMATION:
NAME: WILSON, MARY J.
REGISTRATION NUMBER: 32,955
REFERENCE/DOCKET NUMBER: 620-22
TELECOMMUNICATION INFORMATION:
TELEPHONE: (703) 816-4000
TELEFAX: (703) 816-4100
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 74 amino acids

TYPE: amino acid
STRANDEDNESS: <Unknown>
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 10:
US-10-214-188-10

Query Match 15.4%; Score 203.5; DB 14; Length 74;
Best Local Similarity 58.3%; Pred. No. 1.6e-10; Indels 5; Gaps 3;
Matches 42; Conservative 13; Mismatches 12;

QY 74 VCEKVEAKGRTTYNEVADEIYSELKSM-AHI---GQGFDEKNIRRRVYDAFNVLIALRV 129
DB 4 VCEK-QRKGTTSYNEVADELVAEFSAADNHILPNESAYDQKNIRRRVYDALNVLWAMNII 62

QY 130 AKEKKEIRWML 141
DB 63 SKEKKEIKWIGL 74

Search completed: February 4, 2005, 20:45:29
Job time : 134 secs

AVDPDKRKEKKAAPRTTGWGLREYSKIIVCEKVEAKGRTTYNEVADEIYSELKSNH
 IGOPEKPTRRRVYDAFNVLIALRVIKKEKIRWGLSNVRYEYKIKLEVRKELY
 NKIRNKALLQIEIEKQFDLLONLKLRNLTSSAENVNGIRUPLFVVKTSKARVEIE
 ISDDSHAFHEFNAGAPFDLDDLSILEGVNRTSIEGRAGRTL" #

ORIGIN

Alignment Scores:

Pred. No.:	2,248-97	Length:	1089
Score:	1314.00	Matches:	260
Percent Similarity:	99.62%	Conservative:	0
Best Local Similarity:	99.62%	Mismatches:	1
Query Match:	99.62%	Indels:	0
DB:	6	Gaps:	0

US-10-088-830-2 (1-261) x AX100704 (1-1089)

Qy		1	MetAlaProProArgGlyGlyAlaalaalaaThraLaLeuAsePleuthrGly	20
Db		20	ATGGCGCCTCCCGCGCGAGACTGTGCGGCGGTACCGCGCACTGCACCTCACCGGC	79
Qy		21	ValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGlyAsnAlaVal	40
Db		80	GTCACATTCTCGAACITTCAGTGTCCCCCGCTCCCGAACCGCGCGGTAAATGCCGTC	139
Qy		41	GlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAlaAlaAlaPro	60
Db		140	CNAAGGAAGGGCGTGTTGACC CGCATTAAGATAGGAAGAAGGAGAGGCTGCGGCACCG	199
Qy		61	ArgIleThrGlyTrpGlyLeuArgGluTrpSerLysIleValCysGluLysValGluAla	80
Db		200	AGGATCACCGGTGGGGGCTCCGCGAGTACAGCAAATAAGTTGTGTGAAGAAAGTTGAAGCC	259
Qy		81	LysGlyArgThrThrTyrrAsnGluValAlaAspGluIleTyrrSerGluLeuLysSerMet	100
Db		260	AAAGGAAGAACACATACATCAGGTTGCAGNCGNAATTAATTCAGAGCTGAAGTCCATG	319
Qy		101	AlaHisIleGlyGlnGlyPheAspGluLysAsnIleArgArgArgValTyrrAspAlaPhe	120
Db		320	GCACATATTGGTCAAGGGTTTCATCAGAGAAGAATATTAGCGCGAGAGTGTATGATGCTTTC	379
Qy		121	AsnValIleLeuAlaLeuArgValIleAlaLysGluLysLysGluIleArgTrpMetGly	140
Db		380	AAGTGTCTCATTTGCATCTCGTGTTATTGCAAAAGAAAAAGGAGATACGGTGGATGGGC	439
Qy		141	LeuSerAsnTyrrArgTyrrGluLysIleLysLysLeuGluGluValArgLysGluLeuVal	160
Db		440	CTTTCAAATTCAGATATCAAAAATAAAGAAGCTTGAGGAAGTTTCGTAAAGRACTCGTC	499
Qy		161	AsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeu	180
Db		500	AAACAAGATTAGGAACAAGAAGGCATCTCTCCAGGAATATCGAAAAACAGTTTGATGATCTC	559
Qy		181	GlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsnGlyIle	200
Db		560	CAAAACATCAAGTTACGTAAACCAACACTGGNAAGCTCAGCAGAGAAATGTTAAATGGCATC	619
Qy		201	ArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSer	220
Db		620	CGCCTTCCATTCCGATTGGTTCAAGACATCTAGGAAAGCAAGGGTGGAAAATTGAGATTCA	679
Qy		221	AspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAsp	240
Db		680	GATGACTCCGAAGTTTGCCCATTTCCAGTTCAATGGTGCCACCATTCACATTTGCATGATGAT	739
Qy		241	LeuSerIleLeuGluGlyValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu	260
Db		740	CTCTCAATCTTTAGGGGGTAGGCGTAACAGCATAGGAAGAGCTGCGCGCGCCACCCCTT	799
Qy		261	His 261 800 CAC 802	
Db				

RESULT 2

TSP271917

21001
176717301

LOCUS	TSP2/I91/	1083 bp	mRNA
DEFINITION	Triticum sp.	mRNA for DP protein (dp gene)	

DEFINITION
ACCEPTATION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

REFERENCES

AUTHORS

JOURNAL
OF
MEDICAL
EDUCATION

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

FEATURES

σene

5,177

CDS

3,1118

ORIGIN

Alignment Scores:

Pred. No.:	1.44e-96	Length:	1083
Alignment Scores:		Matches:	258
Score:	1304.00	Conservative:	1
Percent Similarity:	99.23%	Mismatches:	2
Best Local Similarity:	98.85%	Indels:	0
Query Match:	98.86%	Gaps:	0
DB:	8		

US-10-088-830-2 (1-261) x TSP271917 (1-1083)

Ov 1 MetA]apProArG[Vg]vA]aA]aA]aThrA]aA]al.euAspI.euThrGly 20

14 15 16 17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32 33 34 35 36 37 38 39 40 41 42 43 44 45 46 47 48 49 50 51 52 53 54 55 56 57 58 59 60 61 62 63 64 65 66 67 68 69 70 71 72 73 74 75 76 77 78 79 80 81 82 83 84 85 86 87 88 89 90 91 92 93 94 95 96 97 98 99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150 151 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169 170 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188 189 190 191 192 193 194 195 196 197 198 199 200 201 202 203 204 205 206 207 208 209 210 211 212 213 214 215 216 217 218 219 220 221 222 223 224 225 226 227 228 229 230 231 232 233 234 235 236 237 238 239 240 241 242 243 244 245 246 247 248 249 250 251 252 253 254 255 256 257 258 259 260 261 262 263 264 265 266 267 268 269 270 271 272 273 274 275 276 277 278 279 280 281 282 283 284 285 286 287 288 289 290 291 292 293 294 295 296 297 298 299 300 301 302 303 304 305 306 307 308 309 310 311 312 313 314 315 316 317 318 319 320 321 322 323 324 325 326 327 328 329 330 331 332 333 334 335 336 337 338 339 340 341 342 343 344 345 346 347 348 349 350 351 352 353 354 355 356 357 358 359 360 361 362 363 364 365 366 367 368 369 370 371 372 373 374 375 376 377 378 379 380 381 382 383 384 385 386 387 388 389 390 391 392 393 394 395 396 397 398 399 400 401 402 403 404 405 406 407 408 409 410 411 412 413 414 415 416 417 418 419 420 421 422 423 424 425 426 427 428 429 430 431 432 433 434 435 436 437 438 439 440 441 442 443 444 445 446 447 448 449 450 451 452 453 454 455 456 457 458 459 460 461 462 463 464 465 466 467 468 469 470 471 472 473 474 475 476 477 478 479 480 481 482 483 484 485 486 487 488 489 490 491 492 493 494 495 496 497 498 499 500 501 502 503 504 505 506 507 508 509 510 511 512 513 514 515 516 517 518 519 520 521 522 523 524 525 526 527 528 529 530 531 532 533 534 535 536 537 538 539 540 541 542 543 544 545 546 547 548 549 550 551 552 553 554 555 556 557 558 559 560 561 562 563 564 565 566 567 568 569 570 571 572 573 574 575 576 577 578 579 580 581 582 583 584 585 586 587 588 589 590 591 592 593 594 595 596 597 598 599 600 601 602 603 604 605 606 607 608 609 610 611 612 613 614 615 616 617 618 619 620 621 622 623 624 625 626 627 628 629 630 631 632 633 634 635 636 637 638 639 640 641 642 643 644 645 646 647 648 649 650 651 652 653 654 655 656 657 658 659 660 661 662 663 664 665 666 667 668 669 670 671 672 673 674 675 676 677 678 679 680 681 682 683 684 685 686 687 688 689 690 691 692 693 694 695 696 697 698 699 700 701 702 703 704 705 706 707 708 709 710 711 712 713 714 715 716 717 718 719 720 721 722 723 724 725 726 727 728 729 730 731 732 733 734 735 736 737 738 739 740 741 742 743 744 745 746 747 748 749 750 751 752 753 754 755 756 757 758 759 760 761 762 763 764 765 766 767 768 769 770 771 772 773 774 775 776 777 778 779 780 781 782 783 784 785 786 787 788 789 790 791 792 793 794 795 796 797 798 799 800 801 802 803 804 805 806 807 808 809 810 811 812 813 814 815 816 817 818 819 820 821 822 823 824 825 826 827 828 829 830 831 832 833 834 835 836 837 838 839 840 841 842 843 844 845 846 847 848 849 850 851 852 853 854 855 856 857 858 859 860 861 862 863 864 865 866 867 868 869 870 871 872 873 874 875 876 877 878 879 880 881 882 883 884 885 886 887 888 889 890 891 892 893 894 895 896 897 898 899 900 901 902 903 904 905 906 907 908 909 910 911 912 913 914 915 916 917 918 919 920 921 922 923 924 925 926 927 928 929 930 931 932 933 934 935 936 937 938 939 940 941 942 943 944 945 946 947 948 949 950 951 952 953 954 955 956 957 958 959 960 961 962 963 964 965 966 967 968 969 970 971 972 973 974 975 976 977 978 979 980 981 982 983 984 985 986 987 988 989 990 991 992 993 994 995 996 997 998 999 1000 1001 1002 1003 1004 1005 1006 1007 1008 1009 1010 1011 1012 1013 1014 1015 1016 1017 1018 1019 1020 1021 1022 1023 1024 1025 1026 1027 1028 1029 1030 1031 1032 1033 1034 1035 1036 1037 1038 1039 1040 1041 1042 1043 1044 1045 1046 1

1
 2
 3
 4
 5
 6
 7
 8
 9
 10
 11
 12
 13
 14
 15
 16
 17
 18
 19
 20
 21
 22
 23
 24
 25
 26
 27
 28
 29
 30
 31
 32
 33
 34
 35
 36
 37
 38
 39
 40
 41
 42
 43
 44
 45
 46
 47
 48
 49
 50
 51
 52
 53
 54
 55
 56
 57
 58
 59
 60
 61
 62
 63
 64
 65
 66
 67
 68
 69
 70
 71
 72
 73
 74
 75
 76
 77
 78
 79
 80
 81
 82
 83
 84
 85
 86
 87
 88
 89
 90
 91
 92
 93
 94
 95
 96
 97
 98
 99
 100
 101
 102
 103
 104
 105
 106
 107
 108
 109
 110
 111
 112
 113
 114
 115
 116
 117
 118
 119
 120
 121
 122
 123
 124
 125
 126
 127
 128
 129
 130
 131
 132
 133
 134
 135
 136
 137
 138
 139
 140
 141
 142
 143
 144
 145
 146
 147
 148
 149
 150
 151
 152
 153
 154
 155
 156
 157
 158
 159
 160
 161
 162
 163
 164
 165
 166
 167
 168
 169
 170
 171
 172
 173
 174
 175
 176
 177
 178
 179
 180
 181
 182
 183
 184
 185
 186
 187
 188
 189
 190
 191
 192
 193
 194
 195
 196
 197
 198
 199
 200
 201
 202
 203
 204
 205
 206
 207
 208
 209
 210
 211
 212
 213
 214
 215
 216
 217
 218
 219
 220
 221
 222
 223
 224
 225
 226
 227
 228
 229
 230
 231
 232
 233
 234
 235
 236
 237
 238
 239
 240
 241
 242
 243
 244
 245
 246
 247
 248
 249
 250
 251
 252
 253
 254
 255
 256
 257
 258
 259
 260
 261
 262
 263
 264
 265
 266
 267
 268
 269
 270
 271
 272
 273
 274
 275
 276
 277
 278
 279
 280
 281
 282
 283
 284
 285
 286
 287
 288
 289
 290
 291
 292
 293
 294
 295
 296
 297
 298
 299
 300
 301
 302
 303
 304
 305
 306
 307
 308
 309
 310
 311
 312
 313
 314
 315
 316
 317
 318
 319
 320
 321
 322
 323
 324
 325
 326
 327
 328
 329
 330
 331
 332
 333
 334
 335
 336
 337
 338
 339
 340
 341
 342
 343
 344
 345
 346
 347
 348
 349
 350
 351
 352
 353
 354
 355
 356
 357
 358
 359
 360
 361
 362
 363
 364
 365
 366
 367
 368
 369
 370
 371
 372
 373
 374
 375
 376
 377
 378
 379
 380
 381
 382
 383
 384
 385
 386
 387
 388
 389
 390
 391
 392
 393
 394
 395
 396
 397
 398
 399
 400
 401
 402
 403
 404
 405
 406
 407
 408
 409
 410
 411
 412
 413
 414
 415
 416
 417
 418
 419
 420
 421
 422
 423
 424
 425
 426
 427
 428
 429
 430
 431
 432
 433
 434
 435
 436
 437
 438
 439
 440
 441
 442
 443
 444
 445
 446
 447
 448
 449
 450
 451
 452
 453
 454
 455
 456
 457
 458
 459
 460
 461
 462
 463
 464
 465
 466
 467
 468
 469
 470
 471
 472
 473
 474
 475
 476
 477
 478
 479
 480
 481
 482
 483
 484
 485
 486
 487
 488
 489
 490
 491
 492
 493
 494
 495
 496
 497
 498
 499
 500
 501
 502
 503
 504
 505
 506
 507
 508
 509
 510
 511
 512
 513
 514
 515
 516
 517
 518
 519
 520
 521
 522
 523
 524
 525

[illegible]

1	1
2	2
3	3
4	4
5	5
6	6
7	7
8	8
9	9
10	10
11	11
12	12
13	13
14	14
15	15
16	16
17	17
18	18
19	19
20	20
21	21
22	22
23	23
24	24
25	25
26	26
27	27
28	28
29	29
30	30
31	31
32	32
33	33
34	34
35	35
36	36
37	37
38	38
39	39
40	40
41	41
42	42
43	43
44	44
45	45
46	46
47	47
48	48
49	49
50	50
51	51
52	52
53	53
54	54
55	55
56	56
57	57
58	58
59	59
60	60
61	61
62	62
63	63
64	64
65	65
66	66
67	67
68	68
69	69
70	70
71	71
72	72
73	73
74	74
75	75
76	76
77	77
78	78
79	79
80	80
81	81
82	82
83	83
84	84
85	85
86	86
87	87
88	88
89	89
90	90
91	91
92	92
93	93
94	94
95	95
96	96
97	97
98	98
99	99
100	100

DG TGG CAAAGGAAGGGGCCGTGTGACCCCGAATATAAGATAGGAGAAGGAGAGAGGCCTGCCCCCACC

```

194 AGGATCACCGTTGGGGTCCGCGAGTACAGAAAATAGTTCTGAGAAAAGTTGAAGCC 253
QY LysGlyArgThrThrTyrAsnGluValAlaAspGluLeuTyrSerGluLeuLysSerMet 100
Db AAGGAAGAACACATACATAGATGAGTTGCGAGCAATTTATTCAGAGCTGAAGTCCATG 313
QY AlaHisIleGlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPhe 120
Db GCACATATTGGTCAAGGGTTTGATGAGAAGATATTAGCGGAGAGTGATGATGCTTTC 373
QY AsnValLeuIleAlaLeuArgValIleAlaLysGluLysGluLeuArgTyrMetGly 140
Db AACGTTCTCATTCGACTTCGTGTTATTGCAAAAGAAAAAGAGAGATACCGTGGATGGC 433
QY LeuSerAsnTyrArgTyrGluLysIleLysLysLeuGluValValArgLysGluLeuVal 160
Db CTTTCAATTTACAGATATGAAAATAAAGAACCTTGAGGAAGTTCTGTAAGAACTCGTC 493
QY AsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeu 180
Db AACAGATTAGGAACAAAGAGGCACTCTCCAGAAATCGAAAAACAGTTTGATGATCTC 553
QY GlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsnGlyIle 200
Db CAAAACATCAAGTTACGTAAACCAACACACTGGAAGCTCAGCAGAGAATGTTAATGGCATC 613
QY ArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSer 220
Db CGCTTCCATTCGTATGTGTCAGACATCTAGAAAAGCAAGGTGGAATTTGAGATTCA 673
QY AspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAsp 240
Db GATGACTCGAAGTTTCCCATTTTCGAGTTCAATGGTGCCACCATTCACATTCATGATGAT 733
QY LeuSerIleLeuGluGlyValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db CTCTCAATCTTGAGGGGTGAGGGGTAAACAGCATAGGAAGCTGGCGGGCCACCCCTT 793
QY 261 His 261
Db 794 CAC 796

RESULT 3
LOCUS AY224529
DEFINITION Oryza sativa (japonica cultivar-group) isolate 21044 DP protein
ACCESSION AY224529
VERSION AY224529
KEYWORDS mRNA, complete cds.
SOURCE AY224529.1 GI:29367574
ORGANISM Oryza sativa (japonica cultivar-group)
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erbartoideae; Oryzaceae; Oryza.
1 (bases 1 to 885)
Cooper,B., Hutchison,D., Park,S., Guimil,S., Luginbuhl,P.,
Ellero,C., Goff,S.A. and Glazebrook,J.
Identification of rice (Oryza sativa) proteins linked to the
cyclin-mediated regulation of the cell cycle
Plant Mol. Biol. 53 (3), 273-279 (2003)
2311120
1475018
PUBMED
REFERENCE 2 (bases 1 to 885)
AUTHORS Cooper,B., Hutchison,D., Park,S., Guimil,S., Luginbuhl,P.,
Ellero,C., Goff,S.A. and Glazebrook,J.
TITLE Identification of rice (Oryza sativa) proteins linked to the
cyclin-mediated regulation of the cell cycle
JOURNAL Submitted (27-JAN-2003) Torrey Mesa Research Institute, Syngenta
MEDLINE Research and Technology, 3115 Merryfield Row, San Diego, CA 92121,
PUBMED USA
FEATURES
source 1. .885

```

```

/organism="Oryza sativa (japonica cultivar-group)"
/mol_type="mRNA"
/cultivar="Nipponbare"
/isolate="21044"
/db_xref="taxon:39947"
/notes="derived from yeast two-hybrid experiments"
1. .885
/codon_start=1
/product="DP protein"
/protein_id="AA072649.1"
/db_xref="GI:29367575"
/translation="MAPPCCDAAAASAPGLANLIREGAGLPSRPVPPRPCTS
DSFAPISREGDDIPQKVSLSRGGNAAREEGGANGKKTGAQRTGWLGR
EFSKIVSKVEAKGRRTTYNEVADEIFAEKLSITQNGLEFDEKNIRRVYDAFNVLIAI
RVIAKDKEIKWGLTNYVEKIOKLEEVHKELITRIKNKKLQETEKQDCLONIT
LNNQASQRAESVNGILLPLLLIKTSRKARVEIEISDSKFAKPDNGAFPTMDDVYS
ILEAIRNNKRGAGLSIHP"
ORIGIN
Alignment Scores:
Pred. No.: 3 37e-63 Length: 885
Score: 890.50 Matches: 185
Percent Similarity: 73.72% Conservative: 31
Best Local Similarity: 63.14% Mismatches: 36
Query Match: 67.51% Indels: 41
DB: 8 Gaps: 4
US-10-088-830-2 (1-261) x AY224529 (1-885)
QY 1 MetAlaProArgGlyGlyAlaAlaAlaAlaThrAlaAlaLeuAspLeuThrGly 20
Db 1 ATGGCGCCTCCCTCGCGCATGCCGCGGCTGCTCCGCCGCCGCCGCTGGCCAAC 60
QY 21 ValHisIleLeuGluAlaSerSerValProLeuProGlu----- 34
Db 61 CTGCTCATTCGGAAGCGCGCGCTCTCTCCGTCGCGAACGCTACCCGCCATTCGCG 120
QY 34 ----- 34
Db 121 CCCTGTACTTCTGATTCCTTCTCCAAATCTTAGGGAAGGACGATATATCCCCCA 180
QY 35 -----AlaGlyAsnAlaValGlnArgLys----- 43
Db 181 AAAAAATCGTTTCTTACGACGCGCGCGCAATGCGCAGAGAGGAGGAGGC 240
QY 44 GlyAlaValAspProAspLysAspArgLysGluLysAlaAlaProArgIleThr 63
Db 241 GGTGCG-----AACAGGAACCGGAAGAGAGAACCCGGGCGCAGAGGATCACC 291
QY 64 GlyTyrGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAlaLysGlyArg 83
Db 292 GGGTGGGGCTCCGTGAGTTCAGCAAGATAGTTTCTAAGAAAGTTAGGCGCAAGGA 351
QY 84 ThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLysSerMetAlaHisIle 103
Db 352 ACCACATATAATGAGGTTCGCGATGAGATTTTTCGCGAGCTGAAGTCCATTACGCAGAAC 411
QY 104 GlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeu 123
Db 412 GGTCTGGAGTTTTCATGAGAAGATATTAGCGGAGGGTATATGATGCTTTCAATGTGCTC 471
QY 124 IleAlaLeuArgValIleAlaLysGluLysGluIleArgTyrMetGlyLeuSerAsn 143
Db 472 ATTGCAATTCGTGTTATTCAAAAGATAAAGAGATTAAGTGGATGGGCGCTTACTPAT 531
QY 144 TyrArgTyrGluLysIleLysLysLeuGluValArgLysGluLeuValAsnLysIle 163
Db 532 TATAGATACGAAAAGATACAGAAGTTGGAGGAAGTTTCACAAAGAACTCATCACAGGATC 591
QY 164 ArgAsnLysGlyAlaLeuLeuGlnGluIleGluLysGlnPheAspLeuGlnAsnIle 183
Db 592 AAGAATAAGGAAGAGCTTCTCCAGGAATTTGAAAGCAGTTTGTGACCTTCAGATATT 651

```



```

Db 194 AAAAAATCGGTTCTTTTACCGACGGCGCGGCAATCGCGCAGAGGAGGAGGC 253
QY 44 GlyAlaValAspProAspLysAspArgLysLysGluLysAlaAlaProArgIleThr 63
Db 254 GGTGCG-----AACAGGAACGGGAAGAGAGAGACCAGGCGCGCAGAGATCAC 304
QY 64 GlyTrpGlyLeuArgGluTrpSerLysIleValCysGluLysValGluAlaLysGlyArg 83
Db 305 GGGTGGGGCTCGGTGAGTTCAGCAGATAGTTTCTAAGAAAGTTCAGGCGCAAGGA 364
QY 84 ThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHisIle 103
Db 365 ACCACATATAATAGGTTCGCGATGAGATTTTTCGGAGCTGAAGTCCATTACCGACAAC 424
QY 104 GlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeu 123
Db 425 GGTCTGGAGTTTGATGAGAAGATATATTAGCGGAGGGTATATGATGCTTTCATGTGCTC 484
QY 124 IleAlaLeuArgValIleAlaLysGluLysGluIleArgTrpMetGlyLeuSerAsn 143
Db 485 ATTGCAATCGTGTATTGCAAAAGATAAAAAGAGATAAAGTGGATGGGCCCTTACTAAT 544
QY 144 TyrArgTrpGluLysIleLysLysGluLysGluValArgLysGluLeuValAsnLysIle 163
Db 545 TATAGATACGAAAGATACAGAAAGTTGGAGGAAGTTTCACAAAGAACTCATCACCAGATC 604
QY 164 ArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeuGlnAsnIle 183
Db 605 AAGAAATPAGAGAAAGCTTCTCCAGGAATTTGAAGCAGTTTGTACCTTCAGAAATAT 664
QY 184 LysLeuArgAsnGlnThrLeuGluSerSerValAsnValAsnGlyIleArgLeuPro 203
Db 665 ACATTACGCAACCCAGCTAGTCAGAGCCGACGAGAGAAAGTTTAATGGCATCTCTTCG 724
QY 204 PheValLeuValLysThrSerArgLysAlaArgValGluIleGluLysSerAspSer 223
Db 725 TTCTTATTGATCAAGACATCCCGAAAGCAAGGCGTGAATTCAGATTTTCGGAAGATTC 784
QY 224 LysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIle 243
Db 785 AGTTTGCAGGTTTCGACTTCAAGGTGCACCAATTCACCATGATGATGATCAATC 844
QY 244 LeuGluGlyValArgArgAsnSerIleGlyArgAlaGly 256
Db 845 CTTGAAGCCATCAGCGCTAACAAA---GGAAGAGCTGGC 880

RESULT 5
ATH294532 1158 bp mRNA linear PLN 15-DEC-2000
LOCUS Arabidopsis thaliana mRNA for Dp-like protein (dpp gene).
DEFINITION Arabidopsis thaliana
ACCESSION AJ294532
VERSION AJ294532.1 GI:11125652
KEYWORDS DP-like protein; dpp gene.
SOURCE Arabidopsis thaliana
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1
AUTHORS Magyar, Z., Atanassova, A., De Veylder, L., Rombauts, S. and Inze, D.
TITLE Characterization of two distinct DP-related genes from Arabidopsis
thaliana
JOURNAL FEBS Lett. 486 (1), 79-87 (2000)
MEDLINE 20562805
PUBMED 11108847
REFERENCE 2 (bases 1 to 1158)
AUTHORS Magyar, Z.
TITLE Direct Submission
JOURNAL Submitted (06-SEP-2000) Magyar Z., Department Plantgenetica,
Laboratorium voor Genetica, Universiteit Gent, K.L. Ledeganckstraat
35., B-9000, Gent, BELGIUM
FEATURES
Location/Qualifiers
1..1158 source

```

```

/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/variety="Columbia"
/db_xref="taxon:3702"
1..1158
/genes="dpp"
/genes="dpp"
/function="cell cycle regulation"
/codon_start=1
/product="DP-like protein"
/protein_id="CAC15484.1"
/db_xref="GI:11125653"
/db_xref="GOA:O9FNV2"
/translation="MTTGSNSHHHESNNNNNPFSTRSMGTAVSGQS VSTSGSMS
PSRSQTITVTSTSDTTFQRLNLDIOGDFAQSAGSGVKKKKQGRAGRPDKGR
GLRQSMKCEKVESKGRITTYNEVDELVAEFALPNDGTSPOQOYQIRRRVYD
ALNVLAMMDIIISKDKKEIQWRGLPTSLSDIEELKNERLSLRNRIEKKTKAYSQLEEQ
YVGLQNLIOENHLYSSGAPSGVALPFLIVOTRPHATVEVISEDMLVHDFNST
PFELHDNFVLATMKFCDDPOOPNGRNNSQLVCHNPTPENPNKGPSTGPTPQDMYE
THLQSQHQHSQLQIIPMPETNNVTSSADTAPVKSPSLPGINSMKKPEN"
ORIGIN
Alignment Scores: 8.64e-31 Length: 1158
Pred. No.: 492.00 Matches: 104
Score: 66.06% Conservative: 40
Percent Similarity: 47.71% Mismatches: 68
Best Local Similarity: 37.30% Indels: 6
Query Match: 8
DB: 2
US-10-088-830-2 (1-261) x ATH294532 (1-1158)
QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
Db 220 GGTGATGATGCTGCTTCTCAAGGAGCTTCTGGTGTAAAGAAAGAGAGGAGCGGT 279
QY 57 AlaAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGlu 76
Db 280 GCGGCTGGTCCAGATTAAGACTGGAAGAGACTACGTCAATTTAGTATGAAAGTTTGAA 339
QY 77 LysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGlu 96
Db 340 AAGTGGAAAGCAAGCAAGAGCAACTTACAATGAGTTTCACAGCGCTTGTTCGTAA 399
QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 400 TTTGCACATTCCAAATAACGATGGAACATCCCTCGATCAGCAACAGATGATGAGAAAAC 459
QY 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131
Db 460 ATAAGCAGAGAGTATATGCTTTAAACGCTCTCATGGCTATGATATAATATCCAG 519
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
Db 520 GATAAAAAGAAATTCATGGAGAGCTCTCTCCGCGACAGCTTAAGCGCATTTGAAGA 579
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGln 171
Db 580 TTAAGAACAGACAGACTCTCACTTAGACACAGAAATTCAGAAAGAAAACATGCAAT 639
QY 172 GluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeu 191
Db 640 GAACATGGAAGAACAAATATGTAGCCCTTCAGAAATCTGATACAGAGAAATGAGCACT 699
QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210
Db 700 AGCTCAGGAATGCTCCAGTGGCGGTGTCTCTTCTTTATCTCTTCCAGACTCGT 759
QY 211 ArgLysAlaArgValGluIleGluLysSerAspAspSerLysPheAlaHisPheGluPhe 230
Db 760 CCTCACGCAACAGTAGAAGTGGAGATATCAGAAGATATCAGAAGATATGAGCTCGTGCAT 819

```

QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
 Db 820 AACAGCACTCCATTGAGCTCCACGACGACAATTTTGCTCCTCAAGACTATGAAG 873

RESULT 6
 AY086018 1618 bp mRNA linear PLN 14-APR-2003
 LOCUS Arabidopsis thaliana clone 206791 mRNA, complete sequence.
 DEFINITION Arabidopsis thaliana (thale cress)
 ACCESSION AY086018.1 GI:21404728
 VERSION F1.1 CDNA
 SOURCE Arabidopsis thaliana
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi
 1 (bases 1 to 1618)
 Haas,B.J., Volkov,N., Town,C.D., Troukhan,M., Alexandrov,N.,
 Feldmann,K.A., Flavell,R.B., White,O. and Salzberg,S.L.
 Full-length messenger RNA sequences greatly improve genome
 annotation
 Genome Biol. 3 (6), RESEARCH0029 (2002)
 22088475
 MEDLINE 12093376
 PUBMED
 REFERENCE 2 (bases 1 to 1618)
 BROVER,V., TROUKHAN,M., ALEXANDROV,N., LU,Y.-P., FLAVELL,R. and
 FELDMANN,K.
 Full-length cDNA from Arabidopsis thaliana
 Unpublished
 3 (bases 1 to 1618)
 BROVER,V., TROUKHAN,M., ALEXANDROV,N., LU,Y.-P., FLAVELL,R. and
 FELDMANN,K.
 Direct Submission
 Submitted (11-MAR-2002) Ceres, Inc, 3007 Malibu Canyon Road,
 Malibu, CA 90265, USA
 This clone sequence is one of 5,000 Ceres full-length cDNAs made
 available to TIGR and Genbank. The following quality assessment of
 this set was done by comparison with known proteins: two percent of
 the clones are estimated to be 5'-truncated; less than one percent
 are 3'-truncated; approximately two percent represent alternative
 splice variants, including unspliced introns and spliced exons; one
 percent may contain premature stop codons; five percent may have
 frame shifts in a coding region. A sequence is considered to be
 5'-truncated if it lacks the translation initiation start (AUG). A
 sequence is considered to be 3'-truncated if it lacks the
 C-terminal end of the encoded protein. Please note that these cDNA
 sequences are derived from the Wg or Llaer ecotypes and therefore
 may contain polymorphisms when compared to sequences from Col-0.
 GenSet carried out the library production and sequencing of the
 full-length clones. Ceres, Inc. carried out the clustering of the
 5' sequences, selection of clones, and sequence assembly.

FEATURES
 source
 1. 1618
 /organism="Arabidopsis thaliana"
 /mol_type="mRNA"
 /db_xref="taxon:3702"
 /clones="206791"
 249..1406
 /codon_start=1
 /product="unknown"
 /protein_id="AA063227.1"
 /db_xref="GI:21534147"
 /translation="MTTGSNNHHNNHHNNNNNNPSTRGWATVSGQSVSTSSMGS
 PSSRSQETIVTSGSDTTTQLNNLDIQGDAGSGAGVKKKQKQRAAGPKTKR
 GLRQFSKVCESKGRITTYNEVADELVAEFLPNDGTSPOQQQYDKNIIRRVYD
 ALNYLMDIISDKKETOWRLPTSLSDIEELKNERLSLRNRIKKTAYISOLEEQ
 VYGLQNLIRNEHLYSSNAPSGGVALPFIIVOTRPHATVEVEISDMOIVHFDNST
 PFELHDNFVLTWKFCQDPQPPQGNQNSQLVCHNFTPENPKGPTGPTQLDMVE
 THLSQSQHQHQSHQLQIIPMPETNNVTSSADTAPVKPSFLPRIMMSMKPEN"

CDS
 249..1406
 /product="unknown"
 /protein_id="AA063227.1"
 /db_xref="GI:21534147"
 /translation="MTTGSNNHHNNHHNNNNNNPSTRGWATVSGQSVSTSSMGS
 PSSRSQETIVTSGSDTTTQLNNLDIQGDAGSGAGVKKKQKQRAAGPKTKR
 GLRQFSKVCESKGRITTYNEVADELVAEFLPNDGTSPOQQQYDKNIIRRVYD
 ALNYLMDIISDKKETOWRLPTSLSDIEELKNERLSLRNRIKKTAYISOLEEQ
 VYGLQNLIRNEHLYSSNAPSGGVALPFIIVOTRPHATVEVEISDMOIVHFDNST
 PFELHDNFVLTWKFCQDPQPPQGNQNSQLVCHNFTPENPKGPTGPTQLDMVE
 THLSQSQHQHQSHQLQIIPMPETNNVTSSADTAPVKPSFLPRIMMSMKPEN"

ORIGIN
 Alignment Scores:

Pred. No.: 1,31e-30 Length: 1618
 Score: 492.00 Matches: 104
 Percent Similarity: 66.06% Conservative: 40
 Best Local Similarity: 47.71% Mismatches: 68
 Query Match: 37.30% Indels: 6
 DB: 8 Gaps: 2

US-10-088-830-2 (1-261) x AY086018 (1-1618)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
 Db 468 GGTGATGATGCTGGTCTCTCAAGAGAGCTTCTGGTGTAAAGAAGAAGAGGAGCGCGT 527
 QY 57 AlaAlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGlu 76
 Db 528 GCGGCTGCTCCAGATAACACTGGAAGAGAGCTACCTCAATTTAGTATGAAGATTGTGA 587
 QY 77 LysValGluAlaLysGlyArgThrThrTrpAsnGluValAlaAspGluIleTyrSerGlu 96
 Db 588 AAGTGGAAAGCAAGGAAGGACACTTACATGAGGTTCAGACGAGCTTGTCTGAA 647
 QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspClnLysAsn 111
 Db 648 TTTCACCTTCCAAATAACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAC 707
 QY 112 IleArgArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131
 Db 708 ATAAGACGAGAGATATATGATGCTTTAAAGCTTCCTAGGCTATGATATATATATCCAG 767
 QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
 Db 768 GATAAAAGAAATTCATGAGGAGAGCTCTTCCTCGCAAGCTTAAGCCGACATTGAAGAA 827
 QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlu 171
 Db 828 TTAAGAAGCAACGACTCTCCTACTAGGAACAGAAATTGAGAAGAACTCATATATCCAA 887
 QY 172 GluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
 Db 888 GAACTGGAAGACATATAGTATGAGCTTCAGATCTGATACAGAAATGAGCATTATAT 947
 QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210
 Db 948 AGCTCAGGAATGCTCCAGTGGCGGTGTGCTCTCTCTTTTCTTTTCTCCAGACTCGT 1007
 QY 211 ArgLysAlaArgValGluIleSerAspSerLysPheAlaHisPheGluPhe 230
 Db 1008 CTTACGCAACAGTAGAAGTGGAGATATCAGAAGATATGAGCTCGTCTGATTGATTTC 1067
 QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
 Db 1068 AACAGCACTCCATTGAGCTCCACGACGACAATTTTGCTCCTCAAGACTATGAAG 1121

RESULT 7
 AF181998 1640 bp mRNA linear PLN 31-DEC-2003
 LOCUS Populus tremula x Populus tremuloides transcription factor DPl
 DEFINITION (DPl) mRNA, complete cds.
 ACCESSION AF181998
 VERSION AF181998.1 GI:33338105
 KEYWORDS
 SOURCE
 ORGANISM
 Populus tremula x Populus tremuloides
 Populus tremula x Populus tremuloides
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 Rosids; eurosids I; Malpighiales; Salicaceae; Salicaceae; Populus.
 1 (bases 1 to 1640)
 Bhalerao,R.P. and Sandberg,G.
 A DP-1 homolog clone from hybrid Aspen
 Unpublished
 2 (bases 1 to 1640)
 Bhalerao,R.P. and Sandberg,G.
 Direct Submission

Qy	188	GlnThrIeuGluSerSerAlaGluAsnValAsn---GlyIleArgLeuPropheValIeu	208
Db	843	GAACAACTGTACAGCTCAGGAAATGCTCCTAGTGGTGGTGTGCTTTTATTCTG	902
Qy	207	valIysThrSerArgIysAlaArgValGluIleGluIleSerAspSerLysPheAla	226
Db	903	GTCACACACGCCCTCATGCAACTGTTGAAGTGGAGATATCAGAAGATATGCAGCTGGTT	962
Qy	227	HiePheGluPheAsnGlyAlaProPheThrLeuHieAspAspLeuSerIleLeuGluGly	246
Db	963	CACCTTGATTTTAATAGCACTCCCTTCGACTCCATGACGATATTAACGTTCTCAAGCA	1022
Qy	247	ValArg	248
Db	1023	ATGAAA	1028
RESULT	8		
LOCUS	AY307373	1640 bp mRNA linear	PLN 24-JUN-2003
DEFINITION	Populus tremula x Populus tremuloideis Dp1 (Dp1) mRNA, complete cds.		
ACCESSION	AY307373		
VERSION	AY307373.1	GI:32187096	
KEYWORDS	Populus tremula x Populus tremuloideis		
SOURCE	Populus tremula x Populus tremuloideis		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eustosids I; Malpighiales; Salicaceae; Populus.		
REFERENCE	1 (bases 1 to 1640)		
AUTHORS	Espinosa-Ruiz,A., Saxena,S., Schmidt,J., Mellerowicz,E., Bako,L.S. and Bhallerao,R.P.		
TITLE	Differential stage specific regulation of cyclin dependent kinases during cambial dormancy in hybrid aspen		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 1640)		
AUTHORS	Espinosa-Ruiz,A. and Bhallerao,R.P.		
TITLE	Direct Submission		
JOURNAL	Submitted (27-MAY-2003) Department of Forest Genetics and Plant Physiology, Swedish University of Agricultural Sciences (SLU), Umea		
FEATURES	907 34, Sweden		
source	Location/Qualifiers		
	1..1640		
	/organism="Populus tremula x Populus tremuloideis"		
	/mol_type="mRNA"		
	/db_xref="taxon:47664"		
gene	1..1640		
	/genes="Dp1"		
CDS	156..1217		
	/genes="Dp1"		
	/codon_start=1		
	/product="Dp1"		
	/protein_id="AAP73785.1"		
	/db_xref="GI:32187097"		
	/translation="MVAGAGAHLEDDGRHPSSARRGGGAGTTGSWVGQSVSTGSGV GSPSRSEHAMATPADSDTFLRLHLDIADDAATDAANKKKKGQRAVGADKSGR GLQPSIKVCEKVESKGTTYNEVADELVAEPADPSNSYSTPQQQVDEKNIIRRVYD ALNWALMDIISKDKKEIOWKGLPTLSLDIEELKVERLGLRNFKKKAYLQLELEQ FVGLQNLQRLQKSSGNAPSGVSLPILQVTRPHATVEVEISDMQLVHDFNST PELHDDNLYLKAMKFCERPSQDGMAPNPADGEGSSMSYQPOLLASPSPTNTPVR HPTSPPLPGIIRKRVKNEH"		
ORIGIN			
Alignment Scores:			
Pred. No.:	2,39e-29	Length:	1640
Score:	476.50	Matches:	112
Percent Similarity:	59.16%	Conservative:	43
Best Local Similarity:	42.75%	Mismatches:	76
Query Match:	36.13%	Indels:	31
DB:	8	Gaps:	6
US-10-088-830-2 (1-261) x AY307373 (1-1640)			

Db 498 GGATATTATCTAAGGATAAAAGGAAATTCAGTGAAGGGCTTGCTCGG----- 549
 QY 146 rGluLysIleLysLysLeuGluValArgLysGluLeuVal-----AsnLysII 163
 Db 550 -ACAAGTATGAGGATGTTGAAGAAATTAAGACAGAGATCATCGACTGAAGGTAGGAT 608
 QY 163 eArgAsnLysLysAlaLeuLeuGlnGluLysGlnPheAspAspLeuGlnAsnII 183
 Db 609 CGACAAGAAAATGCATATTGTCAGGAGTTAGAAGATCAATTCTAGGTCTTCAAACTT 668
 QY 183 eLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValasn---GlyIleArgLys 202
 Db 669 GCACACGCGAAGCAGAGCTTATGCTTCAGGAAATGCTCTTCAGGAGGAGTGGCAT 728
 QY 202 uProPheValLeuValLysThrSerArgLysAlaArgValGluLysLeuLysSerAspAs 222
 Db 729 GCCATTATATGTCGACGACAGCTCTCATGCTACAGTAGAAGTGGAGATATCAGAAGA 788
 QY 222 pSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSe 242
 Db 789 TATGCAGCTGGTGCATTTTGTATTTCAATAGCACTCCATTTGAACATGCATGACGATTCCT 848
 QY 242 rIleLeuGluGlyValArgAsnSerIleGlyArgAlaGlyArg 257
 Db 849 TGTAAGTAA-----GCATTGGGGTTCCTCGGCAAA 879

RESULT 10
 AX449314
 LOCUS
 DEFINITION Sequence 33 from Patent WO0185946.
 ACCESSION AX449314
 VERSION AX449314.1 GI:21698061

KEYWORDS
 SOURCE Arabidopsis thaliana (thale cress)
 ORGANISM Arabidopsis thaliana
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
 rosids; eustosids II; Brassicales; Brassicaceae; Arabidopsi.

REFERENCE
 AUTHORS Inze,D., Boudolf,V., de Veylder,L., Acosta,J.A. and Magyar,Z.
 TITLE Nucleic acid molecules encoding plant cell cycle proteins and uses
 JOURNAL Patent: WO 0185946-A 33 15-NOV-2001;
 CropsDesign N.V. (BE)

FEATURES
 source
 Location/Qualifiers
 1..1131
 /organism="Arabidopsis thaliana"
 /mol_type="unassigned DNA"
 /db_xref="taxon:3702"

ORIGIN
 Alignment Scores:
 Pred. No.: 9,74e-29 Length: 1131
 Score: 466.50 Matches: 101
 Percent Similarity: 63.76% Conservative: 38
 Best Local Similarity: 46.33% Mismatches: 64
 Query Match: 35.37% Indels: 15
 DB: 6 Gaps: 3

US-10-088-830-2 (1-261) x AX449314 (1-1131)

QY 37 GlyAsnAlaValGlnArgLysGlyValAlaValAspProAspLysAspArgLysLysGluLys 56
 Db 220 GGTGATGATGCTGGTCTCAAGAGCTTCTGGTTAAAGAAAGAGAGGGAGCGCT 279
 QY 57 AlaAlaAlaProArgIleThrGlyLysLeuArgLysSerLysIleValCysGlu 76
 Db 280 GCGGCTGCTCCAGATAGACTGCAAGAGGACTAGCTCAATTAGTATGAAGTTCTGAA 339
 QY 77 LysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluLysLeuTyrSerGlu 96
 Db 340 AAGGTGGAAGCAAGAAAGGACAACTTACATGAGGTTGCAGACGAGCTTGTGTCTGAA 399

QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
 Db 400 TTTGCACTTCCAAATACGATGGAACATCCCTGATCAGCAACGATGATGAGAAAAC 459
 QY 112 IleArgArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131
 Db 460 ATAAGACGAAGAGTATGATGCTTTAAACGCTCCTCATGGCTATGATATATCAAG 519
 QY 132 GluLysLysGluLysLeuArgTyrMetGlyLysSerAsnTyrArgTyrGluLysIleLysLys 151
 Db 520 GATAAAAAAGAAATTCATGGAGAGGCTTCTTCGGCAAGCTTAAGCGCATTTGAAGAA 579
 QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGln 171
 Db 580 TTAAAGAACGAAGCACTCTCATTAGGACAGAAATTCAGAGAAAACCTGATATCCCAA 639
 QY 172 GluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
 Db 640 GAACTGGAAGAACAA-----AGMAATGAGCACTTATAT 672
 QY 192 SerSerAlaGluAsnValasn---GlyIleArgLeuProPheValLeuValLysThrSer 210
 Db 673 AGCTCAGAAAATGCTCCAGTGGCGGTGCTTCTTCCTTTTATCTCTCCAGACTCGT 732
 QY 211 ArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHisPheGluPhe 230
 Db 733 CCTCAGCACACAGTAGAGTGGAGATATCAGAAGATATGCAGCTCGTGCAATTTGATTC 792
 QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
 Db 793 AACAGCACTCCTATTGAGTCCACGACGACAATTTTGTCTCCTCAAGCATATGAAG 846

RESULT 11
 AY224589
 LOCUS
 DEFINITION 1140 bp mRNA linear PLN 05-FEB-2004
 dimerization factor mRNA, complete cds.
 ACCESSION AY224589
 VERSION AY224589.1 GI:29371982

KEYWORDS
 SOURCE Oryza sativa (japonica cultivar-group)
 ORGANISM Oryza sativa (japonica cultivar-group)
 Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Ehrhartoideae; Oryzeae; Oryza.
 1 (bases 1 to 1140)
 Cooper,B., Hutchison,D., Park,S., Guimil,S., Luginbuhl,P.,
 Ellero,C., Goff,S.A. and Glazebrook,J.
 Identification of rice (Oryza sativa) proteins linked to the
 cyclin-mediated regulation of the cell cycle
 Plant Mol. Biol. 53 (3), 273-279 (2003)

JOURNAL MEDLINE
 PUBMED 14750518
 REFERENCE
 2 (bases 1 to 1140)
 Cooper,B., Hutchison,D., Park,S., Guimil,S., Luginbuhl,P.,
 Ellero,C., Goff,S. and Glazebrook,J.
 Direct Submission
 TITLE
 JOURNAL Submitted (27-JAN-2003) Torrey Mesa Research Institute, Syngenta
 Research and Technology, 3115 Merryfield Row, San Diego, CA 92121,
 USA

FEATURES
 source
 Location/Qualifiers
 1..1140
 /organism="Oryza sativa (japonica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="Nipponbare"
 /isolates="31182"
 /db_xref="taxon:39947"
 /note="derived from yeast two-hybrid experiments"
 1..1140
 /codon_start=1
 /product="E2F dimerization factor"
 /protein_id="AA072709.1"
 /db_xref="GI:29371983"

CDS

Db	619	GCAGTAGTCGAATCAGAGATTCTGAGATATGCAACTTGTACACCTCGACTTCATATAGC	678
QY	233	AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245	
Db	679	ACACCTTTCTCGTCCATGATGATCTTACATTTTGAAA 717	
RESULT 14			
LOCUS	ATH319027	960 bp mRNA linear PLN 29-JAN-2002	
DEFINITION	Arabidopsis thaliana partial mRNA for E2F dimerisation partner protein (dp2a gene).		
ACCESSION	AJ319027	1 GI:18447785	
VERSION	AJ319027.1	GI:18447785	
KEYWORDS	dp2a gene; E2F dimerisation partner protein.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.	
AUTHORS	Rosignol, P.		
TITLE	E2F family transcription factors: AtE2F-a and AtDP-a, induce Arabidopsis leaf cells to re-enter S phase		
JOURNAL	Unpublished		
REFERENCE	2	(bases 1 to 960)	
AUTHORS	Bergounioux, C.		
TITLE	Direct Submission		
JOURNAL	Submitted (29-AUG-2001) Bergounioux C., UMR 8618, Cnrs, IBP bat 630		
UNIVERSITE	Paris-Sud, 91405 Orsay, FRANCE		
FEATURES	Location/Qualifiers		
source	1..960	/organism="Arabidopsis thaliana"	
		/mol_type="mRNA"	
		/cultivar="Colombia"	
		/db_xref="taxon:3702"	
		/tissue_type="buds"	
		/dev_stage="flowering"	
	1..960		
gene	/gene="dp2a"		
mRNA	<1..>960		
CDS	79..957		
	/gene="dp2a"		
	/functions="transcription"		
	/codon_start=1		
	/product="E2F dimerisation partner protein (DP2a)"		
	/protein_id="CAC87459.1"		
	/db_xref="GI:18447786"		
	/db_xref="GOA:Q9FNV3"		
	/db_xref="UniProt/TrEMBL:Q9FNV3"		
	/translation="MSMEMELFVTPKQROHVSVEKTPVRKLIIVDDDSIEGSEKK		
	QGRSTGGGLRQFSVMVCQKLEAKITTYKEVADEIISDPATIKQNAEKLNEVNE		
	KNTRRVYDALNFMALDIARQKLEITRWKGLPTCKDVEVMKDRNKMVSSVORKA		
	AFILREKYSLSLSMRNQVMVKTQGPAGFTLPFILLETNPRAVVEIETSDMQ		
	LVLHDSNTPSVHDDAYILKLMQEQKQEQNRVSSSSSTHHQSHSSSSSSSCIAS		
	GTSGPVCWNSGSDTR"		
ORIGIN			
Alignment Scores:			
Pred. No.:	9,48e-26	Length:	960
Score:	428.50	Matches:	88
Percent Similarity:	64.32%	Conservative:	49
Best Local Similarity:	41.31%	Mismatches:	69
Query Match:	32.49%	Indels:	7
DB:	8	Gaps:	2
US-10-088-830-2 (1-261) x ATH319027 (1-960)			
QY	40	ValGlnArgLysGlyAlaValAspProAspLysAspArgLysGlyLysAlaAla 59	
Db	157	GTGAGAGGAATGATTGTTGTTGATGATCTCTGAAATGGATCAGAGAGAAAGGCAA 216	
QY	60	ProArgIleThrGlyTTPGlyLeuArgGluTyrSerLysIleValCysGluLysValGlu 79	
Db	217	TCAAGACTTCGGAGCGGGCTTCGTCATTCAGTGTATGTTGTCAGAGTTGGAA 276	
QY	80	AlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSer 99	
Db	277	GCCAAAGAGATAACTTACTTACAAGGAGGTTGCAGACGAAATATTATTTCAGATTTGCGACA 336	
QY	100	MetalAlaHisIleGlyGln-----GlyPheAspGluLysAsnIleArg 113	
Db	337	ATTAAAGCAAAACGACAGAGAGCGCTTTGAATGAAAATGAGTACAAATGAGAAGAACATAAGG 396	
QY	114	ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133	
Db	397	CGAGAGTCTACATGCGCTCAATGTTTCATGGCGTTGGATATTATTTCAGAGGGATAAA 456	
QY	134	LysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153	
Db	457	AAGGAAATCCGGTGGAAAGGACTTCTATTACTCTCAAAAGGATGTGGAAGAGTCAAG 516	
QY	154	GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIle 173	
Db	517	ATGGATCGTAATAAAGTTAAGACAGTGTGCAAAAGAGCGTCTTTCTTAAAGAGTTG 576	
QY	174	GluLysGlnPheAspAspLeuGlnAsnLysLeuArgAsnGlnThrLeuGluSerSer 193	
Db	577	ACAGAAAGGTTCTCAAGTCTTCAGAGTCTTATGTCAGAAATCAAGAGATGTTGTGAAG 636	
QY	194	AlaGluAsn---ValAsnGlyIleArgLeuProPheValLeuValLysThrSerArgLys 212	
Db	637	ACTCAAGGCCACAGAGAGGATTTTACCTTACCATTCACTTCTACCTGAGACAAACCCCTAC 696	
QY	213	AlaArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPheAsnGly 232	
Db	697	GCAGTAGTCGAATCGAATTCGAGATTCTGAAGATATGCAACTTGTACACCTCGACTTCATAGC 756	
QY	233	AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245	
Db	757	ACACCTTTCTCGTCCATGATGATGCTTACATTTTGAAA 795	
RESULT 15			
LOCUS	AX449297	1114 bp DNA linear PAT 03-JUL-2002	
DEFINITION	Sequence 16 from Patent WO0185946.		
ACCESSION	AX449297		
VERSION	AX449297.1	GI:21698044	
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1	Inze, D., Boudolf, V., de Veylder, L., Acosta, J. A. and Magyar, Z.	
AUTHORS		Nucleic acid molecules encoding plant cell cycle proteins and uses therefor	
TITLE	Patent: WO 0185946-A 16 15-NOV-2001;		
JOURNAL	CropDesign N.V. (BE)		
FEATURES	Location/Qualifiers		
source	1..1114	/organism="Arabidopsis thaliana"	
		/mol_type="unassigned DNA"	
		/db_xref="taxon:3702"	
ORIGIN			
Alignment Scores:			
Pred. No.:	1.14e-25	Length:	1114
Score:	428.50	Matches:	88
Percent Similarity:	64.32%	Conservative:	49
Best Local Similarity:	41.31%	Mismatches:	69
Query Match:	32.49%	Indels:	7
DB:	6	Gaps:	2
US-10-088-830-2 (1-261) x AX449297 (1-1114)			


```
QY 40 ValGlnArgLysGlyAlaValAspProAspLysArgLysLysGluLysAlaAla 59
Db |||:||||| ||||| ||| :||| |||||
92 GTGAGAAGGAATTGATTGATGATTCTGAATTCGATCAGAGAAGAGGCCAA 151
QY 60 ProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGlu 79
Db |||:||||| ||||| |||:||||| |||:||||| |||:||||| |||:|||||
152 TCAAGAATCTTCGAGCGCGCTTCGTCATTCAGTGTATGTTGTCAGAAGTTGGA 211
QY 80 AlaLysGlyArgThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSer 99
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
212 GCCAGAAGATACATCTACAGGAGGTTGCAGACGAATATATTCAGATTTGCCACA 271
QY 100 MetAlaHisIleGlyGln-----GlyPheAspGluLysAsnIleArg 113
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
272 ATTAACCAAAACGCACAGAGCCCTTTGAATGAAATGAGTACATCAGAGAACAATAAG 331
QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
332 CGGAGAGTCTACGATCGCTCAATGTGTTTCATGCGCTTGGATATTATTCAGGGATAAA 391
QY 134 LysGluIleArgTropMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
392 AAGGAATCCGGTGAAGAGACTTCCTATTACTGCAAAAGGATGTGGAAGAAAGTCAAG 451
QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
452 ATGGATCGTAATAAAGTTATGACAGCTGTGCAAAAGAGGCTGCTTTCTTAAAGAGTTG 511
QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
512 AGAGAAAAGGTCTCAAGTCTTGAGAGTCTTATGTCAGAAATCAAGAGATGGTTGTGAAG 571
QY 194 AlaGluAsn--ValAsnGlyIleArgLeuProPheValLeuValLysThrSerArgLys 212
Db ||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
572 ACTCAAGGCCCGCAGCAAGGATTTACTTACCATTATTCTACTTGAGACAACCCCTCAC 631
QY 213 AlaArgValGluIleGluSerAspSerLysPheAlaHisPheGluPheAsnGly 232
Db ||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
632 GCAGTAGTCGAATCGAGATTCTGAGATATGCAACTTGACACCTCGACTCAATAGC 691
QY 233 AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245
Db ||||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
692 ACACCTTTCTCGGTCCATGATGCTTACATTTTGAAA 730
```

Search completed: February 6, 2005, 23:22:05
Job time : 4661 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2005, 18:50:47 ; Search time 583 Seconds
(without alignments)

2650.176 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGAAATAALDLTG.....SILEGVRNSIGRAGRATLH 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DRV=xlh
-Q/cgn2_1/USFTO.spoil/US1008830/runat_04022005_122602_29549/app_query.fasta_1.455
-DB=N Geneseq 16Dec04 -QFMT=fastp -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptio -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=us1008830 @CGN 1.1 470 @runat_04022005_122602_29549 -NCPU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq 16Dec04:.*
1: Geneseqn19808:.*
2: Geneseqn19908:.*
3: Geneseqn20000:.*
4: Geneseqn20010:.*
5: Geneseqn20011:.*
6: Geneseqn20012:.*
7: Geneseqn20013:.*
8: Geneseqn20014:.*
9: Geneseqn20015:.*
10: Geneseqn20016:.*
11: Geneseqn20017:.*
12: Geneseqn20018:.*
13: Geneseqn20019:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1314	99.6	1089	4 AAF80144	Aaf80144 Nucleotid
2	890.5	67.5	885	12 ADO636842	Adq36842 Cell prol
3	492	37.3	1158	12 ADO63082	Ado63082 Transcript
4	492	37.3	1158	12 ADO62142	Ado62142 Transcript
5	492	37.3	1618	3 AAC39795	Aac39795 Arabidops

6	482.5	36.6	1443	12 ADO62852	Ado62852 Transcript
7	476.5	36.1	1153	12 ADO63626	Ado63626 Transcript
8	471.5	35.7	1041	12 ADO36844	Adq36844 Cell prol
9	466.5	35.4	1131	6 AAS96304	Aas96304 Arabidops
10	466.5	35.4	1140	12 ADO36854	Adq36854 Cell prol
11	463	35.1	1245	12 ADO63363	Ado63363 Transcript
12	461	35.0	900	3 AAA59702	Aaa59702 DNA encod
13	456	34.6	1442	6 AAS96332	Aas96332 Arabidops
14	455	34.5	1193	2 AA234580	Aaz34580 Corn DP-2
15	428.5	32.5	1114	6 AAS96287	Aas96287 Arabidops
16	428.5	32.5	1189	6 AAS96319	Aas96319 Arabidops
17	428.5	32.5	1267	12 ADO63084	Ado63084 Transcript
18	428.5	32.5	1267	12 ADO62144	Ado62144 Transcript
19	428.5	32.5	1274	10 ADH59559	Adh59559 Dimerisat
20	428.5	32.5	1274	10 ACC45100	Acc45100 Arabidops
21	399.5	30.3	2320	11 ADI32130	Adi32130 Human cDN
22	399.5	30.3	4359	12 ADN05729	Adn05729 Antipeori
23	396.5	30.1	1396	5 AAS92244	Aas92244 DNA encod
24	394	29.9	1266	6 ABL65866	AbL65866 Lung canc
25	394	29.9	2783	13 ACN38014	Acn38014 Tumour-as
26	394	29.9	2968	5 ABV26838	Abv26838 Human pro
27	394	29.9	2968	5 ABV20992	Abv20992 Human pro
28	389	29.5	1385	2 AAX76730	Aax76730 Murine DP
29	387	29.3	2968	11 ACN88986	Acn88986 Breast ca
30	378.5	28.7	1700	2 AAT10582	Aat10582 Transcript
31	376	28.5	870	6 AAS96410	Aas96410 Arabidops
32	375.5	28.5	2203	4 ABL08071	AbL08071 Drosophil
33	374.5	28.4	692	10 AAD59591	Aad59591 Human tra
34	374.5	28.4	1440	10 ADK67024	Adk67024 Gene #114
35	374.5	28.4	1440	11 ADI32073	Adi32073 Human cDN
36	374.5	28.4	1440	12 ADM78581	Adm78581 Human tra
37	374.5	28.4	1440	12 ADP13337	Adp13337 Renal cel
38	374.5	28.4	1440	13 ADR25496	Adr25496 Breast ca
39	374.5	28.4	2394	12 ADO19868	Ado19868 Human PRO
40	374.5	28.4	2607	4 AAF33503	Aaf33503 Human col
41	372.5	28.2	1700	2 AAG65344	Aag65344 Transcript
42	370	28.1	642	13 ADR65325	Adr65325 Cotton CD
43	361.5	27.4	962	12 ADO63216	Ado63216 Transcript
44	361.5	27.4	2673	6 AAS94979	Aas94979 Human DNA
45	361	27.4	1317	12 ADO62851	Ado62851 Transcript

ALIGNMENTS

RESULT 1
AAF80144
ID AAF80144 standard; DNA; 1089 BP.
XX
AC AAF80144;
XX
DT 11-JUN-2001 (first entry)
XX
DE Nucleotide sequence of a wheat E2F-dimerisation partner (DP) protein.
XX
KW E2F-dimerisation partner; DP protein; E2F transcription factor; G1 phase;
KW S phase; cell cycle; retinoblastoma protein; alter cell proliferation;
KW ss.
XX
OS Triticum monococcum.
XX
FH Key Location/Qualifiers
CDS 20...805
FT /*tag= a
FT /product= "E2F-dimerisation partner (DP) protein"
XX
PN WO200121644-A2.
XX
PD 29-MAR-2001.
XX
PF 25-SEP-2000; 2000WO-EF009325.
XX
PR 24-SEP-1999; 99ES-00002127.
PR 11-NOV-1999; 99ES-00002474.

```
XX (CNSJ ) CONSEJO SUPERIOR INVESTIGACIONES CIENTIF.
PA Gutierrez-Armenta C, Ramirez-Parra E;
PI WPI: 2001-257972/26.
XX P-PSDB; AAB67762.
DR New isolated, enriched, cell free and/or recombinant nucleic acid useful
DR for e.g. altering cell proliferation characteristic such as to alter
XX plant cell, organ or tissue size.
XX Claim 13; Fig 1; 77pp; English.
XX The present sequence encodes a E2F-dimerisation partner (DP) protein. The
CC protein acts as a plant E2F transcription factor. E2F and DP are two
CC proteins that hetero-dimerise to form an active transcription factor that
CC regulates G1 to S phase of the cell cycle, and later, the expression of
CC genes required for S-phase progression. E2F and retinoblastoma protein
CC also interact as a hetero-dimer in cells to suppress certain genes. This
CC repression involves binding of the retinoblastoma protein to the E2F-DP
CC dimer that is in turn bound to sites on DNA through the E2F DNA binding
CC domain. DP proteins can be modulated to alter plant cell, organ or tissue
CC shape, and particularly to alter cell proliferation characteristic such
CC as to alter plant cell, organ or tissue size
XX SQ Sequence 1089 BP; 367 A; 215 C; 261 G; 246 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 4,01e-122 Length: 1089
Score: 1314.00 Matches: 260
Percent Similarity: 99.62% Conservative: 0
Best Local Similarity: 99.62% Mismatches: 1
Query Match: 99.62% Indels: 0
DB: 4 Gaps: 0

US-10-088-830-2 (1-261) x AAF80144 (1-1089)
QY 1 MetAlaProProArgGlyGlyAlaAlaAlaAlaAlaAlaThraAlaLeuLeuLeuThrGly 20
DB 20 ATGGCGCCTCCCGCGCGGAGCTGTGCGCGCTACCGCGCTACCGCGCTGACCGGC 79
QY 21 ValHisLeuLeuGluAlaSerValProLeuProGluAlaGlyGlyAsnAlaVal 40
DB 80 GTGCACATCTCGAAGCTTCCAGTGTCCCGCGCTCCCGCGCTCCCGCGCTAATGCGGTC 139
QY 41 GlnArgLysGlyAlaValAspProAspLysAspArgLysLysGlyLysAlaAlaPro 60
DB 140 CAAAGGAGGGGGCTGTGACCGGATAAGATAGGAGAGGAGAGGCTGCGGACCG 199
QY 61 ArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAla 80
DB 200 AGGATCACCGGTTGGGGGCTCCGCGAGTACAGCAAAATAGTTTGTGAGAAAGTTGAAGCC 259
QY 81 LysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMet 100
DB 260 AAAGGAAGAACACATACATAGAGTTGACAGAGAAATTTATTCAGAGCTGAAGTCAATG 319
QY 101 AlaHisIleGlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPhe 120
DB 320 GCACATATGCTCAAGGGTTTGATGAGAAGAAATATTAGCGGAGAGTGTATGATGCTTTC 379
QY 121 AsnValLeuIleAlaLeuArgValIleAlaLysGluLysLysGluIleArgTrpMetGly 140
DB 380 AACGTTCTCATTCGCTCGTGTATTGCAAAAAAGAAAAAGAGATACGCGTGATGGGC 439
QY 141 LeuSerAsnTyrArgTyrGluLysIleLysLysLeuGluValArgLysGluLeuVal 160
DB 440 CTTTCATATTACAGATATGAAAATTAAGAGCTTGAGGAAGTTCTGTAAGAACTCGTC 499
QY 161 AsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspLeu 180
DB 500 AACAAAGATTAGGAACAAAGAGGCACTCCTCCAGGAAATCGAAAAACAGTTTGTATGATCTC 559
```

```
181 GlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsnGlyIle 200
560 CAAAACATCAAGTTACGTAACCAACACTGGAAGCTCAGCAGAGAAATGTTAATGGCATC 619
201 ArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSer 220
620 CGCCTTCATTCGATTGTTGTCAGACATCTAGGAAGCAAGGTCGAAATTCAGATTTC 679
221 AspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspApp 240
680 GATGACTCGAAGTTTGCCCATTTCCAGTTCAATGTTGTCACCATTCACATTGCATGATGAT 739
241 LeuSerIleLeuGluGlyValArgAsnSerIleGlyArgAlaGlyAlaThrLeu 260
740 CTCTCAATCTTTGAGGGGGTAAGGCGTAACAGCATAGGAAGAGCTGGCCGCCACCCCT 799
261 His 261
800 CAC 802

RESULT 2
ADQ36842
ID ADQ36842 standard; DNA; 885 BP.
XX ADQ36842;
AC ADQ36842;
XX 07-OCT-2004 (first entry)
DE Cell proliferation-related nucleic acid sequence #1.
XX Cell proliferation-related nucleic acid sequence #1.
KW cell proliferation related polypeptide; cell proliferation; senescence;
KW differentiation; stress response; ds.
OS Oryza sativa.
XX WO2004061122-A2.
XX 22-JUL-2004.
XX 23-DEC-2003; 2003WO-US041200.
XX 26-DEC-2002; 2002US-0436565P.
PA (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX Cooper B;
XX WPI; 2004-534388/51.
XX New nucleic acid molecule encoding a cell proliferation-related
PT polypeptide, useful for modulating cell proliferation, senescence,
PT differentiation, development, and stress response in plants, and for
PT producing enhanced food crops.
XX Claim 3; SEQ ID NO 1; 408pp; English.
XX The present invention relates to an isolated nucleic acid molecule
CC encoding a cell proliferation-related polypeptide. The nucleic acid
CC molecule and the encoded polypeptide, and methods are useful for
CC modulating cell proliferation, senescence, differentiation, development,
CC and stress response in plants, and for producing enhanced food crops. The
CC present sequence represents a cell proliferation-related nucleic acid
CC sequence. The present sequence is published separately from the main body
CC of the specification as EPO data.
XX SQ Sequence 885 BP; 249 A; 203 C; 243 G; 190 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 9.32e-80 Length: 885
Score: 890.50 Matches: 185
Percent Similarity: 73.72% Conservative: 31
Best Local Similarity: 63.14% Mismatches: 36
```



```
Alignment Scores: 1,24e-39 Length: 1158
Pred. No.: 492.00 Matches: 104
Score: 66.06% Conservative: 40
Percent Similarity: 47.71% Mismatches: 68
Best Local Similarity: 37.30% Indels: 6
Query Match: 12 Gaps: 2
DB:

US-10-088-830-2 (1-261) x ADO63082 (1-1158)
QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
DB 220 GGTGATGATGCTGTTCTCAAGAGCTTCTGGTGAAGAGAGAGGACAGCT 279
QY 57 AlaAlaAlaProArgLysThrGlyLeuArgLysSerLysLysValCysGlu 76
DB 280 CGCGTGGTCCAGATAAGACTGAAGAGGACTACGTCAATTTAGTATGAAGTTTGTGAA 339
QY 77 LysValGluAlaLysGlyArgThrThrTyraAsnGluValAlaAspGluLysSerGlu 96
DB 340 AAGTGGAAAGAAAGGAAGGACACTTACAATGAGGTTCGACAGAGCTTGTGTGAA 399
QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
DB 400 TTGCACTTCCAAATAACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAC 459
QY 112 IleArgArgValThrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131
DB 460 ATAAGACGAGAGTATATGATGCTTTAAACGCTCTCATGCTATGATATATATATCAAG 519
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyraArgTyraGluLysLysLys 151
DB 520 GATAAAGAAAGAAATTCATGGAGAGGTCTTCCTCGGACAGCTTAAGCGACATGGAAGA 579
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGln 171
DB 580 TTAAGAACGACGACTCTCACTTAGCAAGAGATTGAGAAGAAACTGCATATCCCAA 639
QY 172 GluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
DB 640 GAATCGGAAGAACAAATATGAGCCCTCAGAATCTGATACAGAGAATGAGCACTTAT 699
QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210
DB 700 AGCTCAGAAATGCTCCAGTGGCGGTGTGCTCTTCTTATCTTCTGTCAGACTCGT 759
QY 211 ArgLysAlaArgValGluIleGluLysSerAspSerLysPheAlaHisPheGluPhe 230
DB 760 CCTCAGCAACAGTAGAAGTGGAGATATCAAGAGATATGCGCTCGTGCATTTTGATTTC 819
QY 231 AsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
DB 820 AACGACACTCCATTTGAGTCTCCAGCAGCACAATTTTGTCTCAAGACTATGAAG 873

RESULT 4
ADO62142
ID ADO62142 standard; DNA; 1158 BP.
XX
AC ADO62142;
XX
DT 15-JUL-2004 (first entry)
XX
DE Transcription factor G2981 coding sequence, SEQ ID 609.
XX
KW Plant; transcription factor; transgenic plant; abiotic stress tolerance;
KM osmotic stress tolerance; cold tolerance; heat tolerance;
KW low nitrogen tolerance; low phosphate tolerance; fungal disease;
KW glyphosate resistance; flowering; fertility; seed development; ds.
XX
OS Arabidopsis thaliana.
XX
PN W02004031349-A2.
XX
```

```
PD 15-APR-2004.
XX
PF 18-SEP-2003; 2003WO-US030292.
XX
PR 18-SEP-2002; 2002US-0411837P.
PR 17-DEC-2002; 2002US-0434166P.
PR 24-APR-2003; 2003US-0465809P.
XX
PA (MEND-) MENDEL BIOTECHNOLOGY INC.
XX
PI Jiang C, Heard JB, Ratcliffe O, Creelman RA, Adam LJ, Reuber TL;
PI Riechmann JL, Haake V, Dubell AN, Keddle JS, Sherman BK;
XX
WPI: 2004-330163/30.
DR P-PSDB; ADO62143.
XX
PT New recombinant polynucleotide encoding transcription factor
PT polypeptides, useful for producing transgenic plants with advantageous
PT properties compared to a reference plant.
XX
PS Claim 1; SEQ ID NO 609; 510pp; English.
XX
CC The present invention relates to novel plant transcription factor
CC proteins (I) and nucleotide sequences (II) (ADO61534-ADO63778). The
CC sequences can be used to produce transgenic plants, which overexpress
CC (II), where the transgenic plant has an altered trait as compared to a
CC non-transgenic plant or wild-type plant. The transgenic plant comprises
CC an altered trait selected from increased tolerance to abiotic stress,
CC increased tolerance to osmotic stress, increased tolerance to cold,
CC increased germination in cold, increased tolerance to heat, increased
CC germination in heat, increased tolerance to freezing conditions,
CC increased tolerance to low nitrogen conditions, increased tolerance to
CC low phosphate conditions, increased tolerance to disease, including
CC fungal disease and particularly Erysiphe, Fusarium and Botrytis,
CC increased tolerance to multiple fungal pathogens, increased resistance to
CC glyphosate, increased sensitivity to ABA, reduced sensitivity to ABA,
CC increased sensitivity to ACC, altered sugar sensing, increased tolerance
CC to sugars, altered carbon/nitrogen sensing, early flowering, late
CC flowering, altered flower structure, loss of flower determinacy, reduced
CC fertility, altered shoot meristem development, altered branching pattern,
CC altered stem morphology, altered vascular tissue structure, reduced
CC apical dominance, altered trichome density, altered trichome development,
CC altered trichome structure, altered root development, altered shade
CC avoidance, altered seed development, altered seed ripening, altered seed
CC germination, slow growth, fast growth, altered cell differentiation,
CC altered cell proliferation, altered cell expansion, altered phase change,
CC altered senescence, abnormal embryo development, altered programmed cell
CC death, lethality when overexpressed, altered necrosis patterns, increased
CC plant size, increased biomass, large seedlings, dwarfed plants, dark
CC green leaves, change in leaf shape, increased leaf size and mass, light
CC green or gray leaves, glossy leaves, altered abaxial/adaxial polarity,
CC altered seed coloration, altered seed size, altered seed shape, large
CC seed, increased leaf wax, increased leaf fatty acids, altered seed oil
CC content, altered seed protein content, altered seedprenyl content,
CC altered leaf prenyl lipid content, increased anthocyanin levels, and
CC decreased anthocyanin levels. Note: The sequence data for this patent did
CC not form part of the printed specification, but was obtained in
CC electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1158 BP; 374 A; 261 C; 253 G; 270 T; 0 U; 0 Other;
```

```
Alignment Scores: 1,24e-39 Length: 1158
Pred. No.: 492.00 Matches: 104
Score: 66.06% Conservative: 40
Best Local Similarity: 37.30% Mismatches: 68
Query Match: 12 Gaps: 2
DB:
```

US-10-088-830-2 (1-261) x ADO62142 (1-1158)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56

```
Db 220 GGTGATGATGCTGTTCTCAAGAGCTTCTGGTGTAAAGAGGAGGACGCGT 279
QY 57 AlaAlaAlaProArgIleThrGlyTyrGlyLeuArgGluTyrSerLysIleValCysGlu 76
Db 280 GCGCTGGTCCAGATAGACTGAAGAGGACTACGTCATTTAGTATGAAAGTTTGTA 339
QY 77 LysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGlu 96
Db 340 AAGTGAAGCAAGAGGAGCACTTACATGAGGTTTCACAGAGCTTGTGCTGAA 399
QY 97 LeuLysSerMetalHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 400 TTGTCATCTTCAAATAACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAC 459
QY 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131
Db 460 ATAAGACGAGAGTATGATGCTTTAAACGCTCTCATGCTATGATATATATCCAAG 519
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
Db 520 GATAAAAGAAATTCATGAGAGGCTCTCTCGGCAAGCTTAAGCGACATTTGAAGAA 579
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLysGln 171
Db 580 TTAAGAACCAACGACTCTCACTTAGGAACAGAAATTCAGAAAGAACTGCATATCCCAA 639
QY 172 GluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
Db 640 GAACCTGGAAGAACAAATAGAGCTTCAGAACTGTATACAGAGAATGAGCACTTATAT 699
QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuLysThrSer 210
Db 700 AGCTGAGAAATGCTCCAGTGCCTGGTGTCTTCTTCTTATCTTCTCCAGACTCGT 759
QY 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHisPheGluPhe 230
Db 760 CCTCAGCAACAGTACAGAGTGGAGATATCAGAGATATGAGCTCGTGCATTTTGATTTC 819
QY 231 AsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248
Db 820 AACAGCACTCCATTTGAGCTCCAGCAGCAATTTTCTCCTCAAGACTATGAAG 873

RESULT 5
AAC39795
ID AAC39795 standard; DNA; 1618 BP.
XX
AC AAC39795;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 25934.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN BP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132407P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 19-MAY-1999; 99US-0134768P.
PR 20-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139763P.
PR 22-JUN-1999; 99US-0139817P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
```

```
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144684P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 06-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.

PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Alignment Scores:
Pred. No.: 1.92e-39
Score: 492.00
Percent Similarity: 66.06%
Best Local Similarity: 47.71%
Query Match: 37.30%
DB: 3

US-10-088-830-2 (1-261) x AAC39795 (1-1618)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
Db 468 GGTGATGATGCTGGTCTCAAGAGCTTCTGGTTAAAGAGAGAGAGAGAGAGAGCGT 527
QY 57 AlaAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGlu 76
Db 528 GCGGCTGGTCCAGATAAGACTGGAAGAGGACTAGCTCAATTTAGTATGAAAGTTTGTGAA 587
QY 77 LysValGluAlaLysGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThr 96
Db 588 AAGGTGGAAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 647
QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 648 TTGTCACCTTCCAAATAACGATGGAACATCCCTGATCAGACAGATGATGAGAAAC 707
QY 112 IleArgArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131
Db 708 ATAAGACGAAGAGTATATGATGCTTTAAACGCTCTCATGGCTATGATATATATATCAAG 767
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
Db 768 GATAAAAAAGAAATTCATGAGAGAGCTTCTCTCGCAACAGCTTAAGCGCATTTGAAGAA 827
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGln 171
Db 828 TTAAGAACGACGACTCTCACTTAGAACAGAGATTTAGAGAGAAACTGCATATTTCCCA 887
QY 172 GluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
Db 888 GAACTGGAAGAACATATGTAGGCTTTCAGATCTCAGATCTCAGAGAGAAATGACACTTAT 947
QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210
Db 948 AGCTCAGGAAATGCTCCAGTGGCGGTGTGCTCTTTCCTTTTATCTTTGTCAGACTCGT 1007
```


DR P-PSDB; AAU72514.

XX New cell cycle protein and nucleic acid molecule encoding it useful for

PT regulating cell cycle progression in plants and for identifying

PT modulators which are useful as herbicides or plant growth regulators.

XX

PS Claim 38; Page 185-186; 316pp; English.

XX

CC The invention relates to a novel cell cycle protein (CCP) and the

CC polynucleotides encoding them. CCP is useful for identifying a compound

CC which modulates the activity of the polypeptide and which binds to the

CC polypeptide and an anti-CCP antibody is useful for detecting the presence

CC of CCP in a sample. A CCP modulator is useful for modulating the cell

CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,

CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP

CC nucleic acid and polypeptide molecules are useful as modulating agents in

CC regulating cell cycle progression in plants. CCP is useful to treat

CC disorders characterised by insufficient or excessive production of CCP

CC protein or production of CCP protein forms which have decreased or

CC aberrant activity. Compounds that bind to or modulate the activity of CCP

CC polypeptide are useful as herbicides or plant growth regulators. The

CC polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,

CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,

CC stimulation or enhancement of cell division, DNA replication, seed set,

CC seed size, seed development, tuber, fruit, leaf formation, shoot and root

CC initiation and/or development, module function, dwarfism in plants,

CC senescence, tolerance or resistance to stress. CCP, the polynucleotide

CC and the anti-CCP antibody are useful in agriculture to modulate the

CC protein levels or activity of a protein involved in the cell cycle due to

CC environmental conditions, including abiotic stress such as cold, nutrient

CC deprivation, heat, drought, salt stress, or biotic stress such as

CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate

CC plant architecture, plant quality traits, plant reproduction and seed

CC development, endoreduplication in storage cells, storage tissues and/or

CC storage organs of plants or its parts. CCP is useful as an immunogen to

CC generate antibodies. CCP protein is useful to screen for naturally

CC occurring CCP substrates. The polynucleotide is useful for expressing CCP

CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to

CC modulate CCP activity. The present sequence encodes a CCP protein of the

CC invention

XX

SQ Sequence 1131 BP; 366 A; 256 C; 247 G; 262 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	Length:	Matches:	1131
Score:	4.36e-37	466.50	101
Percent Similarity:	63.76%	Conservative:	38
Best Local Similarity:	46.33%	Mismatches:	64
Query Match:	35.37%	Indels:	15
DB:	6	Gaps:	3

US-10-088-830-2 (1-261) x AAS96304 (1-1131)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56

DB 220 GGTGATGATGCTGGTCTCAGAGGCTTCGGTGAAGAGAGAGAGGGGACAGCGT 279

QY 57 AlaAlaAlaProArgLysThrGlyLeuArgLysLysLysLysValCysGlu 76

DB 280 GCGGCTGGTCCAGATAAGACTGGAAGAGGAGGACTACGTCATTTAGTAAAGTTTGTGAA 339

QY 77 LysValGluAlaLysGlyArgThrThrThrThrThrThrThrThrThrThrThrThrThr 96

DB 340 AAGGTGGAAGCAAGAAAGAGGACAACTTCAATGAGGTTCACACGAGCTTGTGTGAA 399

QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111

DB 400 TTGTCACCTCCAAATACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAC 459

QY 112 IleArgArgValTyrAspAlaPheAsnValLeuLeuAlaLeuArgValIleAlaLys 131

DB 460 ATAAGACGAAGATATGATGCTTTAAACGTCCTCATGCTATGATATATATATATATCAAG 519

132 GluLysLysGluLeuArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLys 151

DB 520 GATAAAGAAGAAATCAATGGAGAGGTCTTCTCGACAGCTTAAGCGACATTAAGAA 579

QY 152 LeuGluGluValArgLysGluLeuValAsnLysLysLysLysLysLysLysLysLys 171

DB 580 TTAAGACGAACGACTCTCCTTAGGAACAGATTGAGAGAAATGCAATATCCCAA 639

QY 172 GluLeuGluLysGlnPheAspAspLeuGlnAsnLysLysLysLysLysLysLysLys 191

DB 640 GAACGTGGAAGAACAA-----AGAAATGAGCAGCTTATAT 672

QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210

DB 673 AGCTCAGGAAATGCTCCAGTGGCGGTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 732

QY 211 ArgLysAlaArgValGluLeuLeuSerAspSerLysPheAlaHisPheGluPhe 230

DB 733 CCTCAGCAACAGTAGAGTGGAGATATCAGAAAGATATGACGCTCGTCATTTGATTTC 792

QY 231 AsnGlyAlaProPheThrLeuHisAspLeuSerLysLysLysLysLysLysLysLys 248

DB 793 AACAGCAGCTCCATTTGAGCTCCACGACGACAAATTTGTCTCAAGACTATGAAG 846

RESULT 10

ADQ36854

ID ADQ36854 standard; DNA; 1140 BP.

XX

AC ADQ36854;

XX

DT 07-OCT-2004 (first entry)

XX

DE Cell proliferation-related nucleic acid sequence #7.

XX

KW cell proliferation related polypeptide; cell proliferation; senescence;

XX

OS Oryza sativa.

XX

PN WO2004061122-A2.

XX

PD 22-JUL-2004.

XX

PF 23-DEC-2003; 2003WO-US041200.

XX

PR 26-DEC-2002; 2002US-0436565P.

XX

PA (SYGN) SYNGENTA PARTICIPATIONS AG.

XX

PI Cooper B;

XX

DR WPI; 2004-534388/51.

XX

PT New nucleic acid molecule encoding a cell proliferation-related

PT polypeptide, useful for modulating cell proliferation, senescence,

PT differentiation, development, and stress response in plants, and for

PT producing enhanced food crops.

XX

PS Claim 3; SEQ ID NO 13; 408pp; English.

XX

CC The present invention relates to an isolated nucleic acid molecule

CC encoding a cell proliferation-related polypeptide. The nucleic acid

CC molecule and the encoded polypeptide, and methods are useful for

CC modulating cell proliferation, senescence, differentiation, development,

CC and stress response in plants, and for producing enhanced food crops. The

CC present sequence represents a cell proliferation-related nucleic acid

CC sequence. The present sequence is published separately from the main body

CC of the specification as EPO data.

XX

SQ Sequence 1140 BP; 327 A; 268 C; 310 G; 235 T; 0 U; 0 Other;

Alignment Scores:

XX SQ Sequence 1245 BP; 414 A; 225 C; 308 G; 298 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,11e-36 Length: 1245
Score: 463.00 Matches: 112
Percent Similarity: 54.80% Conservative: 42
Best Local Similarity: 39.86% Mismatches: 79
Query Match: 35.10% Indels: 48
DB: 12 Gaps: 8

US-10-088-830-2 (1-261) x ADO63363 (1-1245)

```
QY 5 ArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGlyValHisIleLeu 24
DB 4 CGAGGCGACAGCAGCTTCCTCCGCTTGAATATCTGCACATCAACGCG----- 51
QY 25 GluAlaSerSerValProLeuProGluAlaGlyGlyAsnAlaValGlnArgLysGly 44
DB 52 -----GACGACGCGCGCTCGCGAGCTCTACGAGCAAGAAAGAGGAGGC 102
QY 45 -----AlaValAspProAspLysAspArgLysLysGlyLysAlaAlaProArgIle 62
DB 103 ACACGGCGAGTGGTCTCGATAAAGTTAACCGG----- 135
QY 63 ThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLysValGluAlaLysGly 82
DB 136 -----GGACTGCGCCAGTTTAGTATGAAAGTTTGTGAGAAAGTTGAAAGTAAAGG 186
QY 83 ArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHis 102
DB 187 AGAACACATATATAGGTGGCAGATGAACCTTGTGTGAGTTTACAGACCCCAACAT 246
QY 103 -----IleGlyGlnGlyPheAspGluLysAsnIle 112
DB 247 AATATTGAGCCACGATCCTGATACCTTAACCGCAACATATGATGAGAAATATA 306
QY 113 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGlu 132
DB 307 CGACGAGAGCTTTATGATGCTTTGAAATGTTCTGATGCTATGACATTTATCTAAAGT 366
QY 133 LysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeu 152
DB 367 AAAAGGAGATCCAGTGGAGGAGGCTTGCCTG-----ACTAGTATAGTGACATT 417
QY 153 GluGluValArgLysGluLeuVal-----AsnLysIleArgAsnLysLysAlaLeu 169
DB 418 GAAGAAATGAAGACGGAGCTTGTGGGACTGAAAGTAGGATTGAAAGAAAGTGTCTTAT 477
QY 170 LeuGlnGluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsn---Gln 198
DB 478 CTACAGGAGGTACAAGACCAATATGTAGGTCTGCAAAACCTGATCAACGAAATGAACAG 537
QY 189 ThrLeuGluSerSerAlaGluAsnValAsnGlyIleArgLeuProPheValLeuValLys 208
DB 538 TCATATGTTTCAGAAACACGCTTCTGGTGGAGTGGCTTGGCCATTTATCCTAGTTTCA 597
QY 209 ThrSerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHisPhe 228
DB 598 ACCCGACTCATGCTACCTGGAAGTGTGAGATATCAGAGATATGACAGCTGGTGCATTTT 657
QY 229 GluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
DB 658 GACTTCAATAGCACCCCATTTGAGCTGCAGCAGACTCATACAGTCTCTTAAAGAAATGCGA 717
QY 249 -----ArgAsnSerIleGlyArgAlaGlyArgAla 258
DB 718 TTCTGTGAAGAGAACACATGACGACCTCAAGAGTCCGATATCAATGAGGTGAGAGC 777
QY 259 Thr 259
DB 778 TCA 780
```

RESULT 12

AAA59702
ID AAA59702 standard; DNA; 900 BP.

XX AAA59702;
XX 14-NOV-2000 (first entry)

XX DNA encoding a transcription factor designated DP.

XX DP polypeptide; transcription factor; gene transcription; cell cycle;
KW DNA replication; DRTF; differentiation-regulated transcription factor 1;
KW transgenic plant; transformation efficiency; ss.

XX Zea mays.

XX Key Location/Qualifiers

FT 1..900
CDS /tag= a
FT /transl_except= (pos: 1..3, aa: Xaa)
FT /product= "DP polypeptide"
FT /note= "Xaa is any amino acid"

XX WO200047614-A1.

XX 17-AUG-2000.

XX 11-FEB-2000; 2000WO-US003651.

XX 12-FEB-1999; 99US-0119857P.

XX (PION-) PIONEER HI-BRED INT INC.

XX Lowe KS, Gordon-Kamm WJ, Bailey MA, Gregory CA;

XX WPI; 2000-524530/47.

XX P-PSDB; AAB07975.

XX Transgenic plants comprising an expression cassette consisting of a DP
nucleic acid, when expressed, the nucleic acid increases the efficiency
of plant cell transformation by increasing the number of dividing cells
in the plant.

XX Claim 6; Page 70-71; 76pp; English.

XX The present sequence encodes a corn DP (not defined) polypeptide. The
polypeptide is a transcription factor that activates transcription of
numerous genes involved in DNA replication, thus playing a role in the G1
to S transition in the cell cycle. Do, also called DRTF (differentiation-
regulated transcription factor 1) has been shown to form specific
multiprotein complexes with the retinoblastoma susceptibility protein,
p107, cyclins and cdk2. The DP polynucleotide was used to produce a
transgenic plant. The DP nucleic acids and proteins are useful for
increasing transformation efficiency in plants by increasing the number
of dividing cells in the plant (dividing cells may be more receptive to
transformation)

XX Sequence 900 BP; 293 A; 178 C; 231 G; 195 T; 0 U; 3 Other;

Alignment Scores:

Pred. No.: 1.16e-36 Length: 900
Score: 461.00 Matches: 110
Percent Similarity: 55.93% Conservative: 41
Best Local Similarity: 40.74% Mismatches: 71
Query Match: 34.95% Indels: 48
DB: 3 Gaps: 8

US-10-088-830-2 (1-261) x AAA59702 (1-900)

QY 16 LeuAspLeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAla 35

DB 31 CTCGACATCAACGCG-----GACGACGCGCGCTGTCGACGCT 69

QY 36 GlyGlyAsnAlaValGlnArgLysGly-----AlaValAspProAspLysAspArgLys 53
Db CCTACGAGCAAGAAAGAGGAGGACACACGCGAGTGGTCTCTGATAAAGGTAAACCG 129
QY 54 LysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIle 73
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 130
74 ValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIle 93
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 154
154 GTTTGTGAGAAAGTTGAAGTAAGGAGAGAACATATATAGGTGGCAGATGAAC 213
QY 94 TyrSerGluLeuLysSerMetAlaHis-----11e 103
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 214
214 GTTGCTGAGTTTACAGACCCCAACAATAATATTGAGCACCATCTCTGATAACCTAAC 273
QY 104 GlyGlnGlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeu 123
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 274
274 GGCACAATATATGAGAAATAATACGACGAGAGTTTATGATGCTTTGATGTTCTG 333
QY 124 IleAlaLeuArgValIleAlaLysGluLysGluIleArgTrpMetGlyLeuSerAsn 143
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 334
334 ATGGCTATGACATATATCTAAGATAAAGAGATCCAGTGGAGGCTTGGCGGT 393
QY 144 TyrArgTyrGluLysIleLysLysLeuGluValValArgLysGluLeuVal-----160
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 394
394 -----ACTAGTATAGTGACATTGAAGAAATGAGACGGAGCTTGGGACTGAA 444
QY 161 AsnLysIleArgAsnLysAlaLeuLeuGlnGluIleGluLysGlnPheAspLeu 180
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 445
445 GGTAGGATTGAAAGAAAGTGTCTTATACAGGAGCTACAGACCAATATGATGAGTCTG 504
QY 181 GlnAsnIleLysLeuAsn---GlnThrLeuGluSerSerAlaGluAsnValAsnGly 199
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 505
505 CAACCTGATTCACGAATGACATGATGTTGTCAGAAACAGCCCTCTCTGTGGA 564
QY 200 IleArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIle 219
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 565
565 GTGGCTTGCCATTATCTAGTTACAGACCCGACCTCATGCTACCGTGGAGTTGAGATA 624
QY 220 SerAspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAsp 239
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 625
625 TCAGAAGATATGAGCTGGTGTGATTTTGACTTCAATAGCACCCCATTCGAGCTGCACGAC 684
QY 240 AspLeuSerIleLeuGluGlyValArg-----Arg 249
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 685
685 GACTCATAGTCTCAAGAAATGCGGATTTCTGTGGAAGAGAACACATGACGACTCAA 744
QY 250 AsnSerIleGlyArgAlaGlyArgAlaThr 259
Db |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||: 745
745 GAGTCGATATCAAAATCGAGGTGAGACTCA 774

RESULT 13

AAS96332
ID AAS96332 standard; cDNA; 1442 BP.

XX AAS96332;

AC AAS96332;

XX AAS96332;

DT 26-FEB-2002 (first entry)

XX Arabidopsis cDNA encoding cell cycle protein CCP33.

XX Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;

KW plant growth regulator; plant development; abiotic stress; biotic stress;

KW nutrient deprivation; pathogen attack; crop yield.

XX Arabidopsis thaliana.

OS WO200185946-A2.

PN 15-NOV-2001.

XX 15-NOV-2001.

PD 14-MAY-2001; 2001WO-IB001307.

XX 14-MAY-2001; 2001WO-IB001307.

XX 12-MAY-2000; 2000US-0204045P.
PR (CROP-) CROPDESIGN NV.
PA Inze D, Boudolf V, De Veyder L, Acosta JAT, Magyar Z;
PI WPI; 2002-0622249/08.
XX P-PSDB; AAU72542.
DR New cell cycle protein and nucleic acid molecule encoding it useful for
DR regulating cell cycle progression in plants and for identifying
PT modulators which are useful as herbicides or plant growth regulators.
XX Claim 38; Fig 39; 316pp; English.
XX The invention relates to a novel cell cycle protein (CCP) and the
CC polynucleotides encoding them. CCP is useful for identifying a compound
CC which modulates the activity of the polypeptide and which binds to the
CC polypeptide and an anti-CCP antibody is useful for detecting the presence
CC of CCP in a sample. A CCP modulator is useful for modulating the cell
CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP
CC nucleic acid and polypeptide molecules are useful as modulating agents in
CC regulating cell cycle progression in plants. CCP is useful to treat
CC disorders characterised by insufficient or excessive production of CCP
CC protein or production of CCP protein forms which have decreased or
CC aberrant activity. Compounds that bind to or modulate the activity of CCP
CC polypeptide are useful as herbicides or plant growth regulators. The
CC polynucleotide is useful for modifying cell fate, plant development,
CC plant morphology, biochemistry and/or physiology, the length of the G1,
CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
CC stimulation or enhancement of cell division, DNA replication, seed set,
CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
CC initiation and/or development, nodule function, dwarfism in plants,
CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
CC and the anti-CCP antibody are useful in agriculture to modulate the
CC protein levels or activity of a protein involved in the cell cycle due to
CC environmental conditions, including abiotic stress such as cold, nutrient
CC deprivation, heat, drought, salt stress, or biotic stress such as
CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate
CC plant architecture, plant quality traits, plant reproduction and seed
CC development, endoreduplication in storage cells, storage tissues and/or
CC storage organs of plants or its parts. CCP is useful as an immunogen to
CC generate antibodies. CCP protein is useful to screen for naturally
CC occurring CCP substrates. The polynucleotide is useful for expressing CCP
CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to
CC modulate CCP activity. The present sequence encodes a CCP protein of the
CC invention
XX SQ Sequence 1442 BP; 451 A; 321 C; 297 G; 373 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	6.75e-36	Length:	1442
Score:	456.00	Matches:	103
Percent Similarity:	58.54%	Conservative:	41
Best Local Similarity:	41.87%	Mismatches:	68
Query Match:	34.57%	Indels:	34
DB:	6	Gaps:	4

US-10-088-830-2 (1-261) x AAS96332 (1-1442)

QY 37 GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGluLys 56

Db 320 GGTGATGATGCTGGTGTCTCAAGGAGCTCTCTGTGTGAAGAAGAGGAGGACGCT 379

QY 57 AlaAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLys-----72

Db 380 GCGGCTGGTCCAGATAAGACTGGAAGAGACTACGTCATATTAGTATGAAGGCTTATC 439

QY 73 -----IleValCysGluLysValGlu 79

Db 440 TCTTTCTCTGCCCCCTATTATGCTTTTCATCTAAATGCTTTTCAATTGTGAAAGGTGGAA 499

QY 80 AlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSer 99
 Db 500 AGCAAGAGGAGGACAACTTACAATGAGTTGCAGACGAGCTTGTTCGAATTCGCACCT 559
 QY 100 MetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
 Db 560 CCAATATACGATGGAACATCCCTGATCAGCAACAGTATGATGAGAAAACATAGACGA 619
 QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
 Db 620 AGAGTATATGATGCTTTAAACGCTCTCATGGCTATGATATAATATCCAAAGGATAAAAA 679
 QY 135 GluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 154
 Db 680 GAAATTCATGGAGAGGCTCTCCGCAACAGCTTAAGCGCACATTTGAAGAAATTAAGAAC 739
 QY 155 ValArgLysGluLeuValAlaAsnLysIleArgAsnLysLysAlaLeuGluIleGlu 174
 Db 740 GAACGACTCTCATTAGGAACAGAAATGAGAGAAAACCTGCATATCCCAAGACTGGAA 799
 QY 175 LysGlnPheAspAsp-----LeuGlnAsnIle 183
 Db 800 GAACAAGTAAATGAACATCATCGATCTCTCGGCTTATCTGCTTCTCTGCTTCAAGATCTG 859
 QY 184 LysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsn---GlyIleArgLeu 202
 Db 860 ATACAGAGAAATGAGCACTTATATAGCTCAGGAAATGCTCCCAAGTGGCGGTGTGCTCTT 919
 QY 203 ProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSerAspAsp 222
 Db 920 CCTTTATCTTCTCAGACTCTCTCAGCAAGTAGAGTAGAGTAGAGATATCAGAGAT 979
 QY 223 SerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAspLeuSer 242
 Db 980 ATGAGCTCTGTCATTTTGAATTTCAACAGCACTCCATTTGAGCTCCACGACGACAAATTT 1039
 QY 243 IleLeuGluGlyValArg 248
 Db 1040 GTCTCAAGACTATGAAG 1057
 RESULT 14
 ID AAZ34580
 XX AAZ34580 standard; cDNA; 1193 BP.
 AC AAZ34580;
 XX 01-FEB-2000 (first entry)
 XX Corn DP-2 cDNA clone.
 DE DP-2; corn; maize; cell cycle regulatory protein; transcription factor;
 KW herbicide; ss.
 XX Zea mays.
 XX Key Location/Qualifiers
 FH 114..698
 FT /*tag= a
 FT
 XX WO9953075-A2.
 XX 21-OCT-1999.
 XX 08-APR-1999; 99WO-US007638.
 XX 09-APR-1998; 98US-0081132P.
 XX (DUPO) DU PONT DE NEMOURS & CO E I.
 XX Klein TM, Morakinyo IO, Odell JT, Sakai H;
 XX WPI, 1999-633830/54.
 DR

P-PSDB; AAY32164.

Plant-derived cell cycle regulatory proteins.

Claim 12; Page 41; 44pp; English.

This is the nucleotide sequence of a contig assembled from the cDNA inserts in clones p005, cbmfh22r, cdelc.pk001.j13 and cen3n.pk0183.bi. It encodes a portion (see AAY32164) of a corn protein that shows homology to human cell cycle regulatory protein DP-2. The deduced amino acid sequence of this cDNA represents 50% of the middle region of the DP-2 protein. The cDNA clones were isolated from cDNA libraries constructed from corn immature ear, endosperm and developing embryo (20 days after pollination). The invention relates to nucleic acid fragments (see AAZ34575-83) encoding plant CDC-15, DP-1, DP-2 and E2F cell cycle regulatory proteins (see AAY32159-67). It also relates to the construction of a chimeric gene encoding all or a portion of the cell cycle regulatory protein, in sense or antisense orientation, where expression of the chimeric gene results in production of altered levels of the cell cycle regulatory protein in a transformed host cell. The nucleic acids and proteins may be used to facilitate studies of cell cycle regulation in plants, provide genetic tools to enhance cell growth in tissue culture, increase gene transfer efficiency and provide more stable transformations. The proteins may also provide targets to facilitate design and/or identification of cell cycle regulatory proteins that may be useful as herbicides

SQ Sequence 1193 BP; 365 A; 220 C; 305 G; 302 T; 0 U; 1 Other;

Alignment Scores:

Pred. No.:	6,66e-36	Length:	1193
Score:	455.00	Matches:	109
Percent Similarity:	55.93%	Conservative:	42
Best Local Similarity:	40.37%	Mismatches:	71
Query Match:	34.50%	Indels:	48
DB:	2	Gaps:	9

US-10-088-830-2 (1-261) x AAZ34580 (1-1193)

QY	16	LeuAspLeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAla	35
Db	30	CTCGACATCAACGCC-----GACGACGCCGCTGTCGAGGCT	68
QY	36	GlyGlyAsnAlaValGlnArgLysGly-----AlaValAspProAspLysAspArgLys	53
Db	69	CCTACGAGCAAGAGAAAGAGAGGACGACACGCGCAGTGGTCTCTGATAAAGTAACCGG	128
QY	54	LysGluLysAlaAlaProArgIleThrGlyTyrGlyLeuArgGluTyrSerLysIle	73
Db	129	-----GGACTGCCCGCTTTAGTAGTAA	152
QY	74	ValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIle	93
Db	153	GTTTGTGAAAGTTGAAAGTAAAGGAGGAGAACACATATATAGGTGCGAGTGAAT	212
QY	94	TyrSerGluLeuLysSer---MetAlaHisIle-----	103
Db	213	GTTCGTGAGTTTACAGACCCCAACATAATATTGAGGCACCATCTCTGATAACCTAAC	272
QY	104	GlyGlnGlyPheAspGluLysAsnIleArgArgArgValTyrAspAlaPheAsnValLeu	123
Db	273	GCGCAACAATATGATGAGAAAAATATACGAGGCGAGTTTATGATCTTTAATATGTTCTG	332
QY	124	IleAlaLeuArgValIleAlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsn	143
Db	333	ATGGCTATGACATATATATCTAAAGATAAAAGAGATCCAGTGGAGGGCTTCCGCGCG	392
QY	144	TyrArgTyrGluLysIleLysLysLeuGluValValArgLysGluLeuVal-----	160
Db	393	-----ACTAGTATAGTGATATTGAAGAATTTGAAGACTGAGCTTGTGGGACTGAAA	443
QY	161	AsnLysIleArgAsnLysLysAlaLeuGlnGluIleGluLysGlnPheAspArgLeu	180

Db 444 GGTAGATTGAGAGAAAGCTGTTACTTACAGAGCTCAAGATCAATATGATGTTG 503
 QY 181 GlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValaAn--Gly 199
 Db 504 CAAACTGATTCAAGAAATGAGCAATATATGTTTCAGGAACACACCTCTGTGGA 563
 QY 200 IleArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIle 219
 Db 564 GTGGTTTGGCAATCATCTAGTCCAGACCGACCTCATCTACCTGGGAAGTTGAGATA 623
 QY 220 SerAspAspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHisAsp 239
 Db 624 TCAGAGATATGAGCTGGTGCATTTGACTTCAATAGCACCCCAATTCGAGCTGCACGAC 683
 QY 240 AspLeuSerIleLeuGluGlyValArg-----ATG 249
 Db 684 GATCATAGCTCTTAAAGAAATGCGATTTCTGTGGAAGACAAACATGACAGACTCAA 743
 QY 250 AsnSerIleGlyArgAlaGlyAlaThr 259
 Db 744 GAGTCGATATCAAAATGAGGTGAGAGCTCA 773

RESULT 15

AAS96287
 ID AAS96287 standard; cDNA; 1114 BP.

AC AAS96287;

DT 26-FEB-2002 (first entry)

XX Arabidopsis cDNA encoding partial cell cycle protein CCP16.

XX Cell cycle protein; CCP; ss; cell cycle regulation; herbicide;
 KW plant growth regulator; plant development; abiotic stress; biotic stress;
 KW nutrient deprivation; pathogen attack; crop yield.

XX Arabidopsis thaliana.

XX WO200185946-A2.

XX 15-NOV-2001.

XX 14-MAY-2001; 2001WO-1B001307.

XX 12-MAY-2000; 2000US-0204045P.

XX (CROP-) CROPEDESIGN NV.

XX Inze D, Boudolf V, De Veylder L, Acosta JAT, Magyar Z;

XX WPI; 2002-062249/08.

XX P-PSDB; AAU72497.

XX New cell cycle protein and nucleic acid molecule encoding it useful for
 PT regulating cell cycle progression in plants and for identifying
 PT modulators which are useful as herbicides or plant growth regulators.

XX Claim 38; Fig 16; 316pp; English.

XX The invention relates to a novel cell cycle protein (CCP) and the
 CC polynucleotides encoding them. CCP is useful for identifying a compound
 CC which modulates the activity of the polypeptide and which binds to the
 CC polypeptide and an anti-CCP antibody is useful for detecting the presence
 CC of CCP in a sample. A CCP modulator is useful for modulating the cell
 CC cycle or growth of a plant such as Arabidopsis thaliana, rice, wheat,
 CC maize, tomato, alfalfa, oilseed rape, soybean, sunflower and canola. CCP
 CC nucleic acid and polypeptide molecules are useful as modulating agents in
 CC regulating cell cycle progression in plants. CCP is useful to treat
 CC disorders characterised by insufficient or excessive production of CCP
 CC protein or production of CCP protein forms which have decreased or
 CC aberrant activity. Compounds that bind to or modulate the activity of CCP
 CC polypeptide are useful as herbicides or plant growth regulators. The
 CC polynucleotide is useful for modifying cell fate, plant development,

CC plant morphology, biochemistry and/or physiology, the length of the G1,
 CC S, G2 and/or M phase of the cell cycle of a plant, initiation, promotion,
 CC stimulation or enhancement of cell division, DNA replication, seed set,
 CC seed size, seed development, tuber, fruit, leaf formation, shoot and root
 CC initiation and/or development, module function, dwarfism in plants,
 CC senescence, tolerance or resistance to stress. CCP, the polynucleotide
 CC and the anti-CCP antibody are useful in agriculture to modulate the
 CC protein levels or activity of a protein involved in the cell cycle due to
 CC environmental conditions, including abiotic stress such as cold, nutrient
 CC deprivation, heat, drought, salt stress, or biotic stress such as
 CC pathogen attack, to modulate e.g. enhance crop yields, and attenuate
 CC plant architecture, plant quality traits, plant reproduction and seed
 CC development, endoreduplication in storage cells, storage tissues and/or
 CC storage organs of plants or its parts. CCP is useful as an immunogen to
 CC generate antibodies. CCP protein is useful to screen for naturally
 CC occurring CCP substrates. The polynucleotide is useful for expressing CCP
 CC protein, to detect CCP mRNA, or a genetic lesion in a CCP gene and to
 CC modulate CCP activity. The present sequence encodes a CCP protein of the
 CC invention
 XX
 SQ Sequence 1114 BP; 378 A; 197 C; 243 G; 296 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 2,77e-33 Length: 1114
 Score: 428.50 Matches: 88
 Percent Similarity: 64.32% Conservative: 49
 Best Local Similarity: 41.31% Mismatches: 69
 Query Match: 32.49% Indels: 7
 DB: 6 Gaps: 2

US-10-088-830-2 (1-261) x AAS96287 (1-1114)

QY 40 ValGlnArgLysGlyAlaValaAspProAspLysAspArgLysLysGluLysAlaAlaA 59
 Db 92 GTGAGAGAAATGATTGTTGATGATGATCTGAAATTTGGATCGAGAGAAAGGCGAA 151
 QY 60 ProArgIleThrGlyTrpGlyLeuArgGluTrpSerLysIleValCysGluLysValGlu 79
 Db 152 TCAGAGAACTTCGGAGCGGGCTTCGTCATTCAGTTGTTGTTGTTGTTGTTGTTGTTG 211
 QY 80 AlaLysGlyArgThrThrThrAsnGluValAlaAspGluIleTrpSerGluLeuLysSer 99
 Db 212 GCCAAGAGATAACTACTTACAAGGAGGTTGCAGACGAAATTTATTTTCAGATTTTGC 271
 QY 100 MetAlaHisIleGlyCln-----GlyPheAspGluLysAsnIleArg 113
 Db 272 ATTAAGCAAAACGAGAGAGCGCTTTGAATGAAATGAGTACAAATGAGAGAAACATAAG 331
 QY 114 ArgArgValTyAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
 Db 332 CGGAGAGTCTACGATCGGCTCAATGTTTCATCGCGCTTGATATTTATTCAGAGGATAA 391
 QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyArgTyArgLysIleLysLysLeuGlu 153
 Db 392 AAGGAATCCGGTGGAAAGACTTCTTATTCCTGCAAAAGAGATGTGGAAGAAGTCAAG 451
 QY 154 GluValArgLysGluLeuValaAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173
 Db 452 ATGGATCGGTAATAAGTATTGAGCTGTGCAAAAGAGCGCTTTCTTTAAAGAGTTG 511
 QY 174 GluLysGlnPheAspAspLeuGlnAsnIleLysLysLeuArgAsnGlnThrLeuGluSer 193
 Db 512 AGAGAAAAGGTCTCAAGTCTTGAGAGTCTTATCTCGAGAAATCAAGAGATGGTTGTGAG 571
 QY 194 AlaGluAsn---ValAsnGlyIleArgLeuProPheValLeuValLysThrSerArgLys 212
 Db 572 ACTCAAGGCCAGCAGAGAGATTACCTTACCATTCTACTTCTAGACAAACCTCAC 631
 QY 213 AlaArgValGluIleGluLysSerAspSerLysPheAlaHisPheGluPheAsnGly 232
 Db 632 GCAGTAGTCGAATCGAGATTTCTGAAGATATGCAACTTTGTACACCTCGACTTCAATAG 691
 QY 233 AlaProPheThrLeuHisAspLeuSerIleLeuGlu 245

Db 692 ACACCTTCTCGGTCGATGATGCTTACATTTTGAA 730

Search completed: February 6, 2005, 22:04:22
Job time : 588 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2005, 20:19:32 ; Search time 3511 Seconds
(without alignments)
2829.614 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPGGAAAAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table: BLOSUM62

Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame_p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US1008830/runat_04022005_122603_29569/app_query.fasta_1.455
-DB=EST -QFMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1008830 @CNG 1_1 3437 @runat_04022005_122603_29569 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gse1:*
9: gb_gse2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	844.5	64.0	1119	9	CL959517
2	817	61.9	649	6	CD882927
3	694	52.6	612	7	CR289243
4	651	49.4	765	6	CA764996
5	615.5	46.7	607	2	BE456002
6	611.5	46.4	665	6	CA078334
7	575.5	43.6	548	2	BE511883
8	567	43.0	634	7	CF484481
9	505.5	38.3	680	6	CA290160

10	488	37.0	1403	3	CNS09YMA	BX831191 Arabidops
11	477	36.2	1418	3	CNS09YMC	BX831357 Arabidops
12	477	36.2	1568	3	CNS0921R	BX831265 Arabidops
13	467	35.4	702	7	CR286882	CR286882 CR286882
14	466	35.3	992	7	CK278661	CK278661 EST724739
15	463	35.1	1245	3	AY108383	AY108383 Zea mays
16	452	34.3	852	7	CO116984	CO116984 GR_EB019
17	416	31.5	743	7	CN215980	CN215980 29828 Sus
18	413.5	31.3	643	6	C96290	C96290 C96290 Marc
19	406	30.8	550	4	BG040894	BG040894 NXSI_116
20	396.5	30.1	846	1	AL965517	AL965517 AL965517
21	396.5	30.1	1276	3	HSN800731	AL080206 Homo sapi
22	395.5	30.0	990	1	AL547857	AL547857 AL547857
23	394	29.9	583	5	BU579005	BU579005 sar64f10
24	394	29.9	1120	3	CR592926	CR592926 full-leng
25	394	28.9	1737	3	CR597951	CR597951 full-leng
26	393.5	29.8	1098	4	BM560485	BM560485 AGENCOURT
27	393	29.8	2409	3	AK033356	AK033356 Mus muscu
28	393	29.8	2551	3	AK029055	AK029055 Mus muscu
29	388	29.4	2647	3	AK079624	AK079624 Mus muscu
30	384	29.1	2347	3	AK053881	AK053881 Mus muscu
31	380.5	28.8	1070	7	CK230732	CK230732 ILLUMIGEN
32	378.5	28.7	1221	9	AY407516	AY407516 Mus muscu
33	378.5	28.7	2379	3	AK013180	AK013180 Mus muscu
34	378.5	28.7	2459	3	AK019997	AK019997 Mus muscu
35	378	28.7	766	6	CB307985	CB307985 AGENCOURT
36	377	28.6	776	6	BP694534	BP694534 BP694534
37	377	28.6	845	5	BX414733	BX414733 BX414733
38	377	28.6	2343	3	AK044784	AK044784 Mus muscu
39	376.5	28.5	924	7	CK418539	CK418539 AUF_Ipova
40	376	28.5	706	7	CF482425	CF482425 POL1_6 H0
41	374.5	28.4	692	4	BG661040	BG661040 D316 SSH
42	374.5	28.4	784	4	BI856367	BI856367 60387725
43	374.5	28.4	843	4	BI335202	BI335202 60298077
44	374.5	28.4	924	5	BQ945665	BQ945665 AGENCOURT
45	374.5	28.4	1013	5	BM925294	BM925294 AGENCOURT

ALIGNMENTS

RESULT 1

CL959517

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CL959517 1119 bp DNA linear GSS 21-SEP-2004
OsIFCC003026 Oryza sativa Express Library Oryza sativa (indica
cultivar-group) genomic, genomic survey sequence.

CL959517 GI:52373702

GSS.

Oryza sativa (indica cultivar-group)

Oryza sativa (indica cultivar-group)

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

Ehrhartoideae; Oryzaceae; Oryza.

1 (bases 1 to 1119)

Ma, L., Wang, J., Chen, C., Liu, X., Su, N., Li, L., Wang, X., Cao, M.,

Jiao, Y., Sun, N., Zhang, X., Bao, J., Sun, D., Zhao, H., Yuan, L.,

Wong, G.K.S., Deng, X.W. and Wang, J.

An analysis of transcriptional regulation of the rice genome and

its comparison to Arabidopsis

Unpublished (2004)

Contact: Chen Chen

Department of Bioinformatic

Beijing Institute of Genomics

Chinese Academy of Sciences, Beijing 101300, China

Tel: 86-10-80481559

Fax: 86-10-80488676

Email: chenchen@genomics.org.cn

Rice genomic sequence.

Class: exon-trapped.

Location/Qualifiers

1..1119

/organism="Oryza sativa (indica cultivar-group)"

/mol_type="genomic DNA"

Munich Information Center for Protein Sequences

(http://mips.gsf.de)

Plate: 05 row: M column: 03.

Location/Qualifiers

FEATURES

source

1. .765
 /organism="Oryza sativa (indica cultivar-group)"
 /mol_type="mRNA"
 /cultivar="IR64"
 /db_xref="taxon:39946"
 /clone="C0001827"
 /tissue_type="Panicles"
 /dev_stage="Flowering"
 /clone_lib="IRRI Drought Stress Panicle Library"
 /note="Vector: pBluescript II SK+; Water stress was applied by not watering for 4 consecutive days. Panicles were collected from control (well watered) and stressed plants at 2 days before heading, at heading, 50% flowering and 4 days after 50% flowering."

ORIGIN

Alignment Scores:

Pred. No.: 2,15e-62 Length: 765
 Score: 651.00 Matches: 125
 Percent Similarity: 89.57% Conservatives: 21
 Best Local Similarity: 76.69% Mismatches: 17
 Query Match: 49.36% Indels: 0
 DB: 6 Gaps: 0

US-10-088-830-2 (1-261) x CA764996 (1-765)

QY 94 TyrSerGluLeuLysSerMetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113
 DB 3 TTTGGGAGCTGAAGTCCATTCAGCAGACGGTCTGGAGTTTGATGAGAGAAATATTAGG 62
 QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
 DB 63 CGGAGGGTATATGATGCTTTCAATGTGCTCATTTGCAATTCGTGTTATGCAAAAGATAAA 122
 QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
 DB 123 AAGGAGATAAAGTGGATGGGCGCTTACTAATATTATAGATACGAAAGATACAGAAAGTTGGAG 182
 QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173
 DB 183 GAGTTCCACAAAGACTCATCCAGGATCAGAAATAGAAAGCTTCTCCAGGAATTT 242
 QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
 DB 243 GAAAGCAGTTTGTATGACCTTCAGATATATACATTACGCAACAGCGCTAGTCAGAGGCCA 302
 QY 194 AlaGluAsnValAsnGlyIleArgLeuPheValLeuValLysThrSerArgLysAla 213
 DB 303 GCAGAAAGTGTAAATGCACTCCCTTCCTTCTTATTGATCAGAGATCCCGAAGCA 362
 QY 214 ArgValGluIleGluLeuSerAspSerLysPheAlaHisPheGluPheAsnGlyVala 233
 DB 363 AGGTGGAAATGAGATTTCGAGAGATTCAAAGTTTGACCGTTGCACTCAACGGTGCA 422
 QY 234 ProPheThrLeuHisAspLeuSerIleLeuGluGlyValArgArgAsnSerIleGly 253
 DB 423 CCATTCCACATGATGATGATCAATCTTGAAGCCATCAGCGGTACCAACAAAGGA 482
 QY 254 ArgAlaGly 256
 DB 483 AGAGCTGGC 491

RESULT 5

BE456002

LOCUS

DEFINITION

BE456002 607 bp mRNA linear EST 22-OCT-2001
 HVSMeg0019D17f Hordeum vulgare pre-anthesis spike EST library
 cDNA clone HVSMEg0019D17f, mRNA sequence.
 BE456002

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

BE456002.3 GI:16318905

EST.

Hordeum vulgare subsp. vulgare

Hordeum vulgare subsp. vulgare

Eukaryota; Viridiplantae; Stratiotophyta; Embryophyta; Tracheophyta;
 Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 Poideae; Triticeae; Hordeum.

REFERENCE

AUTHORS

1. (bases 1 to 607)

Wing,R., Close,T.J., Kleinhofs,A., Wise,R., Begum,D., Frisch,D.,
 Yu,Y., Henry,D., Palmer,M., Rambo,T., Simmons,J., Choi,D.W.,
 Fenton,R.D., Close,S.J., Oates,R. and Main,D.

TITLE

Development of a genetically and physically anchored EST resource
 for barley genomics: Morex pre-anthesis spike cDNA library

JOURNAL

COMMENT

Unpublished (2001)
 On Jul 26, 2000 this sequence version replaced gi:13155022.

Contact: Wing RA

Clemson University

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total hg bases = 141

Seq primer: AATTAACTCTCACTAAAGG

High quality sequence stop: 585.

FEATURES

source

1. .607

/organism="Hordeum vulgare subsp. vulgare"

/mol_type="mRNA"

/cultivar="Morex"

/sub_species="vulgare"

/db_xref="taxon:112509"

/clone="HVSMEG0019D17f"

/tissue_type="pre-anthesis spike"

/lab_host="SOLR"

/clone_lib="Hordeum vulgare pre-anthesis spike EST library"

/note="Vector: lambdaZAP; Site 1: EcoRI; Site 2: XhoI;

Plants were grown in the greenhouse at the University of
 California, Riverside (Fenton, SJ Close, TJ Close). Whole
 spike with awns trimmed were collected at white, green and
 yellow anther stages (Fenton). Total RNA was prepared from
 each pool, equal quantities of all three RNA pools were
 combined, poly(A) RNA was purified from the mixture, one
 primary unamplified cDNA library was made, and 1 million
 pfu were in vivo excised to give pBluescript SK(-) cDNA
 phagemids. These steps were performed in the TJ Close lab
 (Choi) at the University of California, Riverside.

Phagemids were plated and picked at the Clemson University
 Genomics Institute (CUGI) (Begum, Palmer, Frisch, Atkins
 and Wing) Plasmid DNA preparations, DNA sequencing and
 sequence analysis were performed at CUGI (Wing, Yu,
 Frisch, Henry, Simmons, Oates, Rambo, Main). The sequence
 has been trimmed to remove vector sequence and contains a
 minimum of 100 bases of phred value 20 or above. For more
 details on library preparation and sequence analysis see
 http://www.genome.clemson.edu/projects/barley. To order
 this clone see http://www.genome.clemson.edu/orders Also
 see Close TJ, Wing R, Kleinhofs A, Wise R (2001)

Genetically and physically anchored EST resources for
 barley genomics. Barley Genetics Newsletter 31:29-30.
 (http://wheat.pw.usda.gov/ggpages/bgn/31/cover.html)"

ORIGIN

Alignment Scores:

Pred. No.: 1,51e-58 Length: 607
 Score: 615.50 Matches: 137
 Percent Similarity: 84.30% Conservatives: 8
 Best Local Similarity: 79.65% Mismatches: 22
 Query Match: 46.66% Indels: 5
 DB: 2 Gaps: 1

US-10-088-830-2 (1-261) x BE456002 (1-607)


```

QY 122 valLeulleAlaLeuArgValIleAlaLysGluLysLysGluLleArgTrpMetGlyLeu 141
Db 483 GTCATATCCACTACGTGTATTACAAAAGACCAAAAGAGATAAAGTGGTGGCTTG 542
QY 142 SerAsnTyArgTrpGluLysLysLysLysLysLysLysLysLysLysLysLysLys 161
Db 543 TCTAACTTCCATATGAAAGATA---AACTTGAGGAAGAACTCCCAAGAACTCATGATC 599
QY 162 LysleArgAsnLysLysAlaLeuLeuGln-GluLleGlu-LysGlnPheAspLeuG 181
Db 600 AGGATTAAAGAACAAAGAAAACCTCTCCAGGGAATTTGAACGACACAGTTTGATGAACCTA 659
QY 181 lnAsnLysLysLeuArgAsn 187
Db 660 CAAAATATAAGTTCCGCAAC 679

RESULT 10
CNS09YMA 1403 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTUS672A12 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX831191.1 GI:42454981
VERSION 1
KEYWORDS HTC; GSUT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1403)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
2 (bases 1 to 1403)
Genoscope.
Direct Submission
Submitted (18-NOV-2003) Genoscope - Centre National de Sequencage :
BP 131 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
The sequences are based on single pass reads.
Life Technologies (a division of Invitrogen) members carried out
full-length libraries construction : Temple G.
Genoscope members carried out sequencing and annotation : Castelli
V., Aury J.M., Jaillon O., Wincker P., Menard M., Cruaud C.,
Schachter V., Weissenbach J., Salanoubat M.
URGV INRA : Clepet C., Caboche M.
Annotation is based on the June 2003 version of the Arabidopsis
Genome released by MIPS (Munich Information center for Protein
Sequences). 5 prime and 3 prime are assembled with Phrap.
http://www.genoscope.cns.fr/externe/sequences/Banque_Projet_EF/Full
length
http://www.genoscope.cns.fr/cgi-bin/ggb/ggb?source=Arabidopsis.

FEATURES
source
1..1403
Location/Qualifiers
/organism="Arabidopsis thaliana"
/mol_type="mRNA"
/strain="Col-0"
/db_xref="taxon:3702"
/clone="GSLTUS672A12"
/tissue_type="Adult vegetative tissue"
/plasmid="pCMVSPORT_6"
1..1403
/gene="At5g03415"

ORIGIN
Alignment Scores: 9.68e-44 Length: 1403
Pred. No.: 488.00 Matches: 104
Score:

```

```

Percent Similarity: 65.14% Conservative: 38
Best Local Similarity: 47.71% Mismatches: 70
Query Match: 3 Indels: 6
DB: 3 Gaps: 2

US-10-088-830-2 (1-261) x CNS09YMA (1-1403)

QY 37 GlyAsnAlaValIcInArgLysGlyAlaValAspProAspLysAspArgLysLysGluLys 56
Db 407 GGTGATGATGCTGCTTCTCAAGAGAGCTTCTGGTGTAAAGAAAGAGAGGACGCGT 466
QY 57 AlaAlaAlaProArgIleThrGlyTrpGlyLysLysLysLysLysLysLysLysLys 76
Db 467 GCGGCTGGTCCAGATAAGACTGGAAGAGGACTACGTCATATTTAGTAGATGAGAGTTGTGA 526
QY 77 LysValGluAlaLysGlyArgThrThrTyArgAsnGluValAlaAspGluLleTySerGlu 96
Db 527 AGGTGGAAAGCAAGGAAGGACAACTTACAAATGAGTTGCAGACGAGCTTGTGTCTGAA 586
QY 97 LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 587 TTTGCACCTTCCAAATGACGATGGAACATCCCTGTATCAGCAACAGTATGATGAGAAAAAC 646
QY 112 IleArgArgValTyArgAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLys 131
Db 647 ATAAGACGAAGAGTATATGATGCTTTAAACGTCCTCATGCTGATGATGATATACCCAAAG 706
QY 132 GluLysLysGluLleArgTrpMetGlyLeuSerAsnTyArgTyArgLysLysLysLys 151
Db 707 GATAAAAAGAAATTCATGAGAGGCTCTCTCCGCAAGCTTAAAGCGACATTTGAAGAA 766
QY 152 LeuGluGluValArgLysGluLysGluLysLysLysLysLysLysLysLysLysLys 171
Db 767 TTAAGACGAAGACGACTCTCACTTAGAAGACAGAAATTTGAGAAGAAAACTGCTATATCCCAA 826
QY 172 GluLleGluLysGlnPheAspLysGlnAsnLysLysLysLysLysLysLysLysLys 191
Db 827 GAATCGAAGACAAATATGATGAGCTTCAGAAATCTGATACAGAGAATGAGCACTTATAT 886
QY 192 SerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSer 210
Db 887 AGCTCAGGAATGCTCCAGTGGCGGTGTGCTTCTTCTTTTATCTTGTCCAGACTCGT 946
QY 211 ArgLysAlaArgValGluLleGluLysLysLysLysLysLysLysLysLysLysLys 230
Db 947 CCTCAGCAACAGTACAGTGGAGATATCAGAAGATATCAGCTCGTGTGATTTGATTC 1006
QY 231 AsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248
Db 1007 AACAGCACTCCATTTGAGCTCCAGCAGCAAAATTTGTCTCAAGACTATGAAG 1060

RESULT 11
CNS09YMC 1418 bp mRNA linear HTC 06-FEB-2004
LOCUS Arabidopsis thaliana Full-length cDNA Complete sequence from clone
DEFINITION GSLTUS812A03 of Adult vegetative tissue of strain col-0 of
Arabidopsis thaliana (thale cress).
ACCESSION BX831357.1 GI:42454999
VERSION 1
KEYWORDS HTC; GSUT cDNA.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 1418)
Castelli,V., Aury,J.M., Jaillon,O., Wincker,P., Clepet,C.,
Menard,M., Cruaud,C., Quetier,F., Scarpelli,C., Schachter,V.,
Temple,G., Caboche,M., Weissenbach,J. and Salanoubat,M.
Whole Genome Comparisons and 'Full-length' cDNA Sequences:
A Combined Approach to Evaluate and Improve Arabidopsis Genome
Annotation
Unpublished
JOURNAL

```


DB: 3 Gaps: 2

US-10-088-830-2 (1-261) x CNS09ZIR (1-1568)

QY	37	GlyAsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGluLys	56
Db	450	GGTGATGATGCTGGTTCACAGGAGCTTTCTGGTGTAAAGAAGAAGAGGGACACGC	509
QY	57	AlaAlaAlaProArgIleThrGlyTTPGLyLeuArgGluTyrSerLysIleValCysGlu	76
Db	510	CGCGCTGGTCCAGTAAGACTGGAGAGGACTACGCAATTAGTAGAGATTGTGA	569
QY	77	LysValGluAlaLysGlyArgThrThrTyAsnGluValAlaAspGluIleTyrSerGlu	96
Db	570	AAGGTGMAAACAAGGAGGACAACTTCAATGAGGTTGCAGACGAGCTTGTGTCTGAA	629
QY	97	LeuLysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn	111
Db	630	TTTGCACTTCCAATTACGATGGAACATCCCCGTATCAGCAACGATGATGAGAAAAC	689
QY	112	IleArgArgArgValTyrAspAlaPheAsnValLeulleAlaLeuArgValIleAlaLys	131
Db	690	ATAAGACGAGAGATATATGATGCTTTAAACGCTCATGGCTATGGATATAATATCCAAG	749
QY	132	GluLysLysGluIleArgThrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys	151
Db	750	GATAAAAAAGAAATTCATGAGAGGCTTCTCCGGAACAAGCTTAAGCACATTTGAAGAA	809
QY	152	LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGln	171
Db	810	TTAAGAACCAACGACTCTCACTTAGGAACAGAAATTCAGAGAACAACCTGCATATTCCAA	869
QY	172	GluIleGluLysGlnPheAspLeuGlnAsnIleLysLeu-ArgAsnGlnThrLeuGln	191
Db	870	GAATCGAAGAACATATGATAGGCCITCAGAAUCTTGATACAGAGAATGAGACATTATA	929
QY	191	uSerSerAlaGluAsnValAsn---GlyIleArgLeuProPheValLeuValLysThrSe	210
Db	930	TAGCTCAGGAAATGCTCCCAGTGGCGGTGTGCTCTTCCTTATCCTTGTCCAGACTCG	989
QY	210	rArgLysAlaArgValGluIleSerAspAspSerLysPheAlaHisPheGluPh	230
Db	990	TCCTCAGCAACAGTAGAAGTGGAGATATCAGAAGATATGACGCTCGTCATTTGATTT	1049
QY	230	eAnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg	248
Db	1050	CACAGCACTCCATTGAGCTCCACGACGACAATTTGTCTCTCAAGACTATGAAG	1104

CR286882 702 bp mRNA linear EST 27-FEB-2004

CR286882 Oryza sativa library (Han B) Oryza sativa cDNA clone p730d0lp5, mRNA sequence.

ACCESSION CR286882

VERSION CR286882.1 GI:44673448

KEYWORDS EST.

SOURCE Oryza sativa

ORGANISM Oryza sativa

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoidae; Oryzeae; Oryza.

AUTHORS Han,B., Feng,Q., Huang,Y.C., Ying,K., Li,Y., Guan,J.P., Zhu,J.J., Zhao,Q., Hu,X., Liu,Y.L., Wu,J., Yu,Z., Chen,L., Fan,D.L., Wang,Q.J., Zhang,L., Lu,Y.Q., Yu,S.L., Liu,X.H., Lu,T.T., Zhang,Y.J., Lu,Y., Li,C., Li,T., Zhang,Y., Hu,H., Jia,P.X., Zhang,L., Lan,L.F., Chen,W., Wu,S.A. and Xue,Y.B.

TITLE Rice cDNA EST clone

JOURNAL Unpublished (2003)

COMMENT Contact: Han Bin
National Center for Gene Research
Chinese Academy of Sciences
500# Cao Bao Road, Shanghai 200233, China

ORIGIN

Alignment Scores:

Pred. No.:	5.25e-41	Length:	1245
Score:	463.00	Matches:	112
Percent Similarity:	54.80%	Conservative:	42
Best Local Similarity:	39.86%	Mismatches:	79
Query Match:	35.10%	Indels:	48
DB:	3	Gaps:	8

US-10-088-830-2 (1-261) x AY108383 (1-1245)

Qy	5	ArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuAspLeuThrGlyValHisIleLeu	24
Db	4	CGAGGCGACACACAGCTTCCTCCGCTTGATAATTCGACATCAACGGC-----	51
Qy	25	GluAlaSerSerValProLeuProGluAlaGlyAsnAlaValGlnArgLysGly	44
Db	52	-----GACGACGGCGCTCGTCGCGAGCTCTACGAGCAAGAGAAAGGAGAGGC	102
Qy	45	-----AlaValAspProAspLysAspArgLysLysGlyLysAlaAlaProArgIle	62
Db	103	ACACGGCGAGTGGTCTGATAAAGTAACCGG-----	135
Qy	63	ThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGlyLysValGluAlaLysGly	82
Db	136	-----CGACTGCGCCAGTTTAGTATGAAAGTTTGTGAGAAAGTTGAAAGTAAAGGG	186
Qy	83	ArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeuLysSerMetAlaHis	102
Db	187	AGAACACATATAATAGGTGGCAGATGAACCTTGTGCTGAGTTTACAGACCCCAACAT	246
Qy	103	-----IleGlyGlnGlyPheAspGluLysAsnIle	112
Db	247	AATATTGAGGCACCATCTCTGATACCCCTTAACGCGCAACAATATGATGAGAAATATA	306
Qy	113	ArgArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGlu	132
Db	307	CGACGAAGAGCTTTATGATGCTTTGAATGTTCTGATGGCTATGGACATTTATCTAAAGAT	366
Qy	133	LysLysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeu	152
Db	367	AAAAAGAGATCCAGTGAAGGGCTTGCCCGCT-----ACTAGTATAAGTGACATT	417
Qy	153	GluGluValArgLysGluLeuVal-----AsnLysIleArgAsnLysLysAlaLeu	169
Db	418	GAAGAATGAAGACGGAGCTTGTGGACTGAAGGTAGGATTGAAAGAAAGTGCTTAT	477
Qy	170	LeuGlnGluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsn---Gln	188
Db	478	CTACAGGAGCTACAAGACCAATATGTAGGTCTGCAAAACCTGATTCACGAAATGAACAG	537
Qy	189	ThrLeuGluSerSerAlaGluAsnValAsnGlyIleArgLeuProPheValLeuValLys	208
Db	538	TCATATGTTTCAGAAACACGCCCTTCTGTGGAGTGGCTTTGCCATTTATCTTAGTTCAG	597
Qy	209	ThrSerArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHisPhe	228
Db	598	ACCGACCTCATGCTACCGTGGAGAGTTGAGATATCAGAAGATATGACGTGGTGCAATTT	657
Qy	229	GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg	248
Db	658	GACTTCAATAGCACCCCATTCGAGCTGCACGACCTCATACGCTCCTTAAAGAAATGCGA	717
Qy	249	-----ArgAsnSerIleGlyArgAlaGlyArgAla	258
Db	718	TTCTGTGAAGAGAACCAACATGACAGCACTCAAGAGTCGATATCAAAATGGAGGTGAGGC	777
Qy	259	Thr	259
Db	778	TCA	780

Search completed: February 7, 2005, 00:20:47
Job time : 3518 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2005, 21:42:57 ; Search time 196 Seconds
(without alignments)

2178.920 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPPGGAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/usfto.spool/US1008830/runat_04022005.122603.29577/app_query.fasta_1.455
-DB=Issued Patents NA -OFMT=fastap -SUFFIX=p2n.rni -MINMATCH=0.1 -LOOPEL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US1008830 @CGN_1.1.59 @runat_04022005.122603.29577 -NCPU=6 -ICPU=3
-NO_MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:

1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	399.5	30.3	2320	4	US-09-023-655-1456
2	389.5	29.5	1157	2	US-08-723-415B-5
3	389.5	29.5	1157	3	US-09-189-627A-5
4	389.5	29.5	1157	3	US-09-710-861-5
5	389	29.5	1202	2	US-08-723-415B-7
6	389	29.5	1202	3	US-09-189-627A-7
7	389	29.5	1202	3	US-09-710-861-7
8	389	29.5	1385	2	US-08-723-415B-1
9	389	29.5	1385	3	US-09-189-627A-1
10	389	29.5	1385	3	US-09-710-861-1
11	385	29.2	1427	4	US-09-949-016-3349
12	384	29.1	1154	2	US-08-723-415B-3

13	384	29.1	1154	3	US-09-189-627A-3	Sequence 3, Appli
14	384	29.1	1154	3	US-09-710-861-3	Sequence 3, Appli
15	378.5	28.7	1700	2	US-08-428-131-1	Sequence 1, Appli
16	378.5	28.7	1700	2	US-08-602-846-1	Sequence 1, Appli
17	378.5	28.7	1700	3	US-09-078-596-1	Sequence 1, Appli
18	374.5	28.4	1440	4	US-09-023-655-1399	Sequence 1399, Ap
19	374.5	28.4	1672	4	US-09-949-016-2937	Sequence 2937, Ap
20	261.5	19.8	358	4	US-09-640-211A-566	Sequence 566, App
21	254.5	19.3	364	4	US-09-640-211A-465	Sequence 465, App
22	187.5	14.2	352	4	US-09-640-211A-440	Sequence 440, App
C 23	156	11.8	601	4	US-09-949-016-120497	Sequence 120497,
C 24	156	11.8	601	4	US-09-949-016-120498	Sequence 120498,
25	156	11.8	193169	4	US-09-949-016-15091	Sequence 15091, A
26	155.5	11.8	58273	4	US-09-949-016-14679	Sequence 14679, A
27	155	11.8	601	4	US-09-949-016-105247	Sequence 105247,
28	146.5	11.1	271	4	US-09-640-211A-576	Sequence 576, App
29	128.5	9.7	1314	3	US-09-242-737-1	Sequence 1, Appli
30	128.5	9.7	2456	1	US-07-882-711-1	Sequence 1, Appli
31	128.5	9.7	2456	2	US-08-462-174-1	Sequence 1, Appli
32	128.5	9.7	2457	2	US-08-723-415B-12	Sequence 12, Appli
33	128.5	9.7	2457	3	US-09-189-627A-12	Sequence 12, Appli
34	128.5	9.7	2457	3	US-09-710-861-12	Sequence 12, Appli
35	128.5	9.7	2486	4	US-09-949-016-4380	Sequence 4380, Ap
36	128.5	9.7	2492	1	US-08-139-937-13	Sequence 13, Appli
37	128.5	9.7	2492	4	US-09-023-655-1230	Sequence 1230, Ap
38	128.5	9.7	2492	5	PCT-US93-11310-13	Sequence 13, Appli
39	128.5	9.7	2517	1	US-08-306-691B-18	Sequence 18, Appli
40	128.5	9.7	2517	1	US-08-385-142-2	Sequence 2, Appli
41	128.5	9.7	2517	2	US-08-481-814A-1	Sequence 1, Appli
42	128.5	9.7	2517	3	US-08-801-092-2	Sequence 2, Appli
43	128.5	9.7	2517	3	US-09-517-584A-3	Sequence 3, Appli
44	128.5	9.7	2517	3	US-09-315-113-2	Sequence 2, Appli
45	123.5	9.4	2078	4	US-09-799-451-486	Sequence 486, App

ALIGNMENTS

RESULT 1

US-09-023-655-1456
; Sequence 1456, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESS: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:

```
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1456:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2320 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: 9604478
; US-09-023-655-1456

Alignment Scores:
Pred. No.: 6.43e-39 Length: 2320
Score: 399.50 Matches: 94
Percent Similarity: 58.33% Conservative: 53
Best Local Similarity: 37.30% Mismatches: 88
Query Match: 30.29% Indels: 17
DB: 4 Gaps: 7

US-10-088-830-2 (1-261) x US-09-023-655-1456 (1-2320)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
Db 222 GTTACTCAGACACATAGCAGAACTACTGGCTGGTCCCT-----GGT 266
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 267 GATAGAAAACGGCTAGAAAATTATAGACTCTGATTTTTCAGAAAGTAAACGACAA 326
QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 327 AAAGGAGATAAA---AATGGGAAGGCTTGAGACACTTTTCAATGAAGTGTGTGAGAA 383
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 384 GTTCACGAAAGGTACACATCGTACATCAATCAAGTCTGATGAGCTGGTGCAGAGTTC 443
QY 98 LysSer-----MetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113
Db 444 ACCAATTCAAATAACCAATTGGCTGCTGATTGTCAGGCTTATGATCAGAAAGACATTAGG 503
QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db 504 CGAAGAGTTTATGCTTTAAATGTCTTAATGCTTAATGCAATGAACATAATTTCAAAGGAAAA 563
QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db 564 AAAGAAATCAAGTGGATGGCTGCTGCCTACCAATTGCTCAGGAATGTCAGAACTCGGAG 623
QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIle 173
Db 624 ATAGAGAAGCAGAGCGGATAGAACGATAAAGCAGAGCGGCCCGCAGCTGCAAGAACTT 683
QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db 684 CTCCTACAGCAAAATCGCTTTCAAACCACTGGTGCAGAGAAATCGCAAAATGACGACAA 743
QY 194 AlaGluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThr 209
Db 744 AACAGGCGCGCGGCTCTGAACCTTACCATTCAGCTGCCATTCATAATCATCAATACA 803
QY 210 SerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---Phe 228
Db 804 AGCAGAAAAACAGTCATATGCGCATCTCCAGTGAC---AAGTTTGATGATCTCTTTC 860
QY 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGlyValArg 248
Db 861 AATTTTGACAAACACC---TTTGAGATCCATGATGACATAGATGACTAAAGCGGATGGGA 917
QY 249 ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 917
```

```
Db 918 ATGTCGTTGGCTGGAGTCAGCAAAATGCTCTCTG 953

RESULT 2
US-08-723-415B-5
; Sequence 5, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: Lathangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1157 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1110
; US-08-723-415B-5

Alignment Scores:
Pred. No.: 3.73e-38 Length: 1157
Score: 389.50 Matches: 94
Percent Similarity: 57.54% Conservative: 51
Best Local Similarity: 37.30% Mismatches: 74
Query Match: 29.53% Indels: 33
DB: 2 Gaps: 8

US-10-088-830-2 (1-261) x US-08-723-415B-5 (1-1157)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCACATAGCTGAGGCTGCT----- 111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 112 -----GCTGGGTTCACGTAACGAAAGCAAAAGGAGATAAA--- 150
QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGGAAGGCTTGAGACATTTTTCATGAAGGTGTGTGAGAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
```


[illegible]

US-10-088-830-2 (1-261) x US-09-189-627A-5 (1-1157)

Qy 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACCTCAGACTCACATAGCTGAGGCTGCT----- 111
Qy 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGluLysAla 57
Db 112 -----GGCTGGTTCACAGTAACGACGAACAAAAAGAGATAA--- 150
Qy 58 AlaAlaProArgIleThrGlyTyrGluLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGAAAGGCTTGAGACATTTTCAATGAAGGTGTGTAGAAA 195
Qy 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 196 GTTCAGCGGAAGGCACAACTTCATACAATGAGGTAGCTGATGAGTCTGATCTGAGTTT 255
Qy 98 LysSer-----MetalHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113
Db 256 ACCAACTCAAATAACCATCTGGCAGCTGATTCGCAGGCTTATGATCAGGAGCAATTAGA 315
Qy 114 ArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLysGluLys 133
Db 316 CGAAGAGTTTATGATGCTTTAAATGTACTAATGGCGATGAACAATAATTTCAAAGGAAAA 375
Qy 134 LysGluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db 376 AAGAATCAAGTGGATTGGCTGCCTACCAATTCGCTCAGGAATGCCAGAACCTGGA 435
Qy 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGlu 173
Db 436 ATCAGAAGCAGAGGCGGATAGAACGGATAAAGCAGAAGCAGGCCAGCTTACAAGAACTT 495
Qy 174 GluLysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db 496 CTCCTTCAGAAATTCGTTTTAAACAACTCGTACAGAGAAATCGACAAAATGAACACAA 555
Qy 194 AlaGluAsn-----ValAsnGly-----IleArgLeuProPheValLeuValLysThr 209
Db 556 AACCAGGGCCCTCCAGCTGTGAATCCACCATTCAGCTGCCATTTATATCAATTAATACA 615
Qy 210 SerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---Phe 228
Db 616 AGCAGGAAAAACAGTCATAGACTGCAGCATCTCCAGTGAC---AAATTGAATACCTTTT 672
Qy 229 GluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
Db 673 AATTTTGATAACACC---TTTGGATCCAGCAGACATAGAGGTACTGAAGCGATGGGA 729
Qy 249 ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 730 ATGCTCTTGGTCTGGAGTCAGGCAATGCTCTCTG 765

RESULT 4
US-09-710-861-5
; Sequence 5, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susan
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; CURRENT FILING DATE: 2000-11-13
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0

```
; SEQ ID NO 5
; LENGTH: 1157
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1110)
US-09-710-861-5

Alignment Scores:
Pred. No.: 3,73e-38 Length: 1157
Score: 389.50 Matches: 94
Percent Similarity: 57.54% Conservative: 51
Best Local Similarity: 37.30% Mismatches: 74
Query Match: 29.53% Indels: 33
DB: 3 Gaps: 8

US-10-088-830-2 (1-261) x US-09-710-861-5 (1-1157)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGATGCTGAGGCTGCT-----111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 112 -----GGCTGGTCCAGTAACGAAGCAAAAGAGAGATAA---150
QY 58 AlaAlaProArgIleThrGlyTyrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGAAAGGCTTGAGACATTTTCAATGAAGGTGTGTGAGAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 196 GTTCAGCGGAAGGCAACATTCATACATGAGTAGCTGATGAGCTGGTATCTGAGTTT 255
QY 98 LysSer-----MetAlaHisIleGlyGlnGlyPheAspGluLysAsnIleArg 113
Db 256 ACCAACTCAATAACCATCTGGCAGCTGATTCGCAGGCTTATGATCAGGAGACATTAGA 315
QY 114 ArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLys 133
Db 316 CGAAGAGTTTATGATGCTTTAAATGTACTAATGCGATGAACATAATTTCAAAGGAAAA 375
QY 134 LysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 153
Db 376 AAGAATCAAGTGGATTGGCTTCCCTGCTACCAATTCCTCAGGAATGCCAGAACCTGGAA 435
QY 154 GluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIle 173
Db 436 ATCGAAGACGAGCGGATAGAACGGATAGAAGCAGAGCGAGCCAGCTACAGAACTT 495
QY 174 GluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSer 193
Db 496 CTCCTTCAGCAAAATGCTTTTAAAAACCTGGTACAGAGAAATCGCAAAATGAACAACA 555
QY 194 AlaGluAsn-----ValAsnGly-----IleArgLeuProPheValLeuValLysThr 209
Db 556 AACCAGGCGCCCTCAGCTGGAATTCACCATTCAGCTGCATTTATATATCATATATACA 615
QY 210 SerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---Phe 228
Db 616 AGCAGGAAACAGTCATAGACTCAGCATCTCCAGTGAC---AAATTTGATACCTTTT 672
QY 229 GluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluValArg 248
Db 673 AATTTTGAATACACC---TTTGATCCACGACATAGAGGTACTGAGCGGATGGGA 729
QY 249 ArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 730 ATGTCCTTTGGTCTGAGTCAAGCAATGCTCTCTG 765

US-08-723-415B-7

; Sequence 7, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: LaThangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1202 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1155
US-08-723-415B-7

Alignment Scores:
Pred. No.: 4,57e-38 Length: 1202
Score: 389.00 Matches: 98
Percent Similarity: 57.48% Conservative: 48
Best Local Similarity: 38.58% Mismatches: 86
Query Match: 29.49% Indels: 22
DB: 2 Gaps: 10

US-10-088-830-2 (1-261) x US-08-723-415B-7 (1-1202)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGATGCTGAGGCT-----GCTGGCTGG 117
QY 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54
Db 118 GTTCCAGTGATAGAAAACAGCAGTAGAGAAATTTATAGACTCTGATTTTTCAGAAAGTAA 177
QY 55 GluLysAlaAlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleVal 74
Db 178 CGAAGCAAAAGGAGAGATAA---AATGGAAAGGCTTGAGACATTTTCAATGAAGGTG 234
QY 75 CysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyr 94
Db 235 TGTGAAAGTTTCAGCGGAAAGGCAACTTCATACATGAGTAGCTGATGAGCTGGTA 294
QY 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
```



```
Alignment Scores:
Pred. No.: 5,7e-38 Length: 1385
Score: 389.00 Matches: 98
Percent Similarity: 57.48% Conservative: 48
Best Local Similarity: 38.58% Mismatches: 86
Query Match: 29.49% Indels: 22
DB: 3 Gaps: 10

US-10-088-830-2 (1-261) x US-09-710-861-1 (1-1385)
QY 18 LeuThGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 265 GTCACTCAGACTCAGACTGAGGCT-----GCTGGCTGG 300
QY 38 AsnAlaValGlnArgLysGlyAla-----ValAspProAspLysAspArgLysLys 54
Db 301 GTTCCAGTAGTAGAAAACGAGCTAGAGATTTATAGACTCTCATTTTTCAGAAAGTAA 360
QY 55 GluLysAlaAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleVal 74
Db 361 CGAAGCAAAAAGAGAGATAAA---AATGGGAAAGGCTTGAGACATTTTTCATGAAGGTG 417
QY 75 CysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyr 94
Db 418 TGTGAGAAAGTTTCAGCGGAAGGACACACTTCATACATAGGTAGCTGATGAGCTGGTA 477
QY 95 SerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsn 111
Db 478 TCTGAGTTTACCAACTCAATAACCATCTGGCAGCTGATTCGGCTTATGATCAGGAGAAC 537
QY 112 IleArgArgValTyrAspAlaPheAsnValIleAlaLeuArgValIleAlaLys 131
Db 538 ATTAGACGAGAGTTTATGATGCTTTAAATGTACTAATGGCGATGACATATTTTCAAAG 597
QY 132 GluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLys 151
Db 598 GAAAAAAGAAATCAAGTGGATGGCTGCTCCCTACCAATTCCTGCTCAGGATGCCAGAAC 657
QY 152 LeuGluGluValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGluGln 171
Db 658 CTGGAATTCGAGAAGCAGAGCGCGATAGAACGGATAAGCAGAGCGAGCCAGCTACAA 717
QY 172 GluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGlu 191
Db 718 GAATCTTCTCTCAGCAAAATGCTTTTAAAAACCTGTGTACAGAGAAATCGCAAAATGAA 777
QY 192 SerSerAlaGluAsn-----ValAsnGly---IleArgLeuProPheValLeuVal 207
Db 778 CAACAAACACGCGCCCTCAGCTGTGATTCACCAATTCAGCTGCATTTATATCAAT 837
QY 208 LysThrSerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis 227
Db 838 AATACAAGCAGGAAACAGCTCATAGACTGCAGCATCTCCAGTGCAC---AAATTTGAATAC 894
QY 228 ---PheGluPheAsnGlyAlaProPheThrLeuHisAspLeuSerIleLeuGluGly 246
Db 895 CTTTTTAATTTTATAACACC---TTTGAGATCCAGCAGACATAGAGTACTGAAGCGG 951
QY 247 ValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 952 ATGGGAATGTCCTTTGGTGGAGTCAGGCAATGTCTCTG 993

RESULT 11
US-09-949-016-3349
; Sequence 3349, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
```

```
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3349
; LENGTH: 1427
; TYPE: DNA
; ORGANISM: Human
; US-09-949-016-3349
```

```
Alignment Scores:
Pred. No.: 1,87e-37 Length: 1427
Score: 385.00 Matches: 92
Percent Similarity: 58.82% Conservative: 48
Best Local Similarity: 38.66% Mismatches: 84
Query Match: 29.19% Indels: 14
DB: 4 Gaps: 8
```

US-10-088-830-2 (1-261) x US-09-949-016-3349 (1-1427)

```
QY 34 GluAlaGlyGlyAsnAlaValGlnArgLysGlyAla-----ValAspProAspLys 50
Db 336 GAATCTCAGTCCAAACAAAGTGTAGAAAACGGGCTAGAAAATTTTATAGACTCTGATTT 395
QY 51 AspArgLysLysLysLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyr 70
Db 396 TCAGAAAGTAAACGAAGCAAAAAGAGAGATAAA---AATGGGAAAGGCTTGAGACACTTT 452
QY 71 SerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAla 90
Db 453 TCAATGAAAGTGTGTGAGAAAGTTCAACGAAAGGTACAACATCGTACAATGAAGTCGCT 512
QY 91 AspGluIleTyrSerGluLeu---LysSerMetAlaHisIleGly-----GlnGlyPhe 107
Db 513 GATGAGCTGGTGTGACAGATTCACCAATTCAAATAACCATTTGGCTGCTGATTCGGCTTAT 572
QY 108 AspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAlaLeuArg 127
Db 573 GATCAGAAACACATTAGCGCAAGAGTTTATGATGCTTTAAATGTGCTAATGGCAATGAAC 632
QY 128 ValIleAlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGlu 147
Db 633 ATAATTTCAAAGGAAAAAAGAAATCAAGTGGATGGCTGCTGCTACCAATTCCTGCTCAG 692
QY 148 LysIleLysLysLeuGluValArgLysGluLeuValAsnLysIleArgAsnLysLys 167
Db 693 GAATGTGAGATCTGGAGATAGAGAGAGCGGATAGAACGGGATAAGACGAGACGCG 752
QY 168 AlaLeuLeuGluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsn 187
Db 753 GCCCAGCTGCAGAACTTCTCTACAGCAAAATCGCTTTCAAAAACCTGGTACAGAGAAAT 812
QY 188 GlnThrLeuGluSerSerAlaGluAsnVal-----AsnGlyIleArgLeuPro 203
Db 813 CGCAAAATGAGCAGCAAAAACCGCGGCTCTGAACTCTCACTACCATTCAGTGCCTGCA 872
QY 204 PheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSerAspAspSer 223
Db 873 TTCATAATCATCAATACAGCAAAAACAGTCATAGATTGGCAGCATCTCCAGTGAC--- 929
QY 224 LysPheAlaHis---PheGluPheAsnGlyAlaProPheThrLeuHisAspLeuSer 242
Db 930 AGTTTGTAGTATCTTTTCAATTTTGACACACC---TTTGAGATTCATGATGACATAGAA 986
QY 243 IleLeuGluGlyValArgArgAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 987 GTACTAAAGCGGATGGATGTGCTTTGGCTGCTGAGTCAGGCAAAATGCTCTCTG 1040
```

RESULT 12

US-08-723-415B-3
; Sequence 3, Application US/08723415B
; Patent No. 5859199
; GENERAL INFORMATION:
; APPLICANT: LaThangue, Nicholas B.
; APPLICANT: delaluna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS
; TITLE OF INVENTION: THEREOF
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: NIXON & VANDERHYE P.C.
; STREET: 1100 No. 5859199th Glebe Rd. 8th floor
; CITY: Arlington
; STATE: VA
; COUNTRY: USA
; ZIP: 22201-4741
; COMPUTER READABLE FORM: disk
; MEDIUM TYPE: Floppy
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/723,415B
; FILING DATE: 30-SEP-1996
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: GB 9610195.1
; FILING DATE: 15-MAY-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Crawford, Arthur R.
; REGISTRATION NUMBER: 25,327
; REFERENCE/DOCKET NUMBER: 117-220
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 703-816-4000
; TELEFAX: 703-816-4100
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1154 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1107
US-08-723-415B-3

Alignment Scores:
Pred. No.: 1154
Score: 384.00
Length: 1154
Percent Similarity: 57.37%
Conservative: 50
Best Local Similarity: 37.45%
Mismatches: 75
Query Match: 29.11%
Indels: 32
Gaps: 9
DB:

US-10-088-830-2 (1-261) x US-08-723-415B-3 (1-1154)

QY 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCACATAGCTGAGGCTGCT-----111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysGluLysAla 57
Db 112 -----GCTCGGGTCCAGTAAACGAGCAAAAGGAGATAAA---150
QY 58 AlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGGAAGGCTTGAGCATTTTTCATGAGGTGTGTGAGAA 195
QY 78 ValGluAlaLysGlyArgThrThrThrThrThrThrThrThrThrThrThrThr 97
Db 196 GTTCAGCGGAAGGCAACTTCATCAATAGGTAGCTGATGAGCTGCTGATCTGAGTTT 255

QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 256 ACCAACTCAATAACCACTGCGCAGCTGATTGGCTTATGATCAGAGAACATTAGACGA 315
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 316 AGAGTTTATGATGCTTTAAATGCTACTAATGGCGATGAACATAATTTCAAAGGAAAAAAA 375
QY 135 GluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 154
Db 376 GAAATCAAGTGAATGGCTGCTTACCAATCTGCTCAGAAATGCCAGAACCTGGAAATC 435
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGlu 174
Db 436 GAGAAGCAGAGCGCGGATAGACGGAATAAGCAGAGAGCGGAGCCAGCTACAGAACTTC 495
QY 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAla 194
Db 496 CTTCAAGCAAAATGCTTTTAAAAACCTGGTACAGAGAAATCGACAAATGAACAAAC 555
QY 195 GluAsn-----ValAsnGly---IleArgLeuProPheValLeuValLysThrSer 210
Db 556 CAGGCGCTCCAGCTGTGAATTCACCATTCAGCTGCCATTTATATCAATTAATCAAGC 615
QY 211 ArgLysAlaArgValGluIleSerAspArgSerLysPheAlaHis---PheGlu 229
Db 616 AGGAAACAGTCATAGACTGCGAGCATCTCCAGTGAC--AAATTTGAATACCTTTTAA 672
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 249
Db 673 TTTGATAACACC--TTTGAGATCCAGCAGACATAGAGGTACTGAAGCGGATGGAATG 729
QY 250 AsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 730 TCCTTTGGTCTCGAGTCAGGCAAAATGCTCTCTG 762

RESULT 13

US-09-189-627A-3
; Sequence 3, Application US/09189627A
; Patent No. 6159691
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; APPLICANT: de la Luna, Susana
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/189,627A
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1154
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1107)
US-09-189-627A-3

Alignment Scores:
Pred. No.: 1154
Score: 384.00
Length: 1154
Percent Similarity: 57.37%
Conservative: 50
Best Local Similarity: 37.45%
Mismatches: 75
Query Match: 29.11%
Indels: 32
Gaps: 9
DB:

US-10-088-830-2 (1-261) x US-09-189-627A-3 (1-1154)

```
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGACTGAGCTGCT-----111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 112 -----GGCTGGTTCAGTAACGAGCAAGCAAAAGGAGATAA---150
QY 58 AlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGAAAGGCTTGAGACATTTTCAATGAAGGTGTGTGAGAAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 196 GTTCAGCGGAAGGCAACATTCATCAATGAGTAGCTGAGCTGGTATCTGAGTTT 255
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 256 ACCAACTCAAATAACCATCTGGCAGCTGATTTCGGCTTATGATCAGGAGAACATTAGACGA 315
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 316 AGAGTTTATGATGCTTTAAATGCTTAATGCGATGAACATAATTTCAAAGGAAAAA 375
QY 135 GluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 154
Db 376 GAATCAAGTGGATGGCTGCTACCAATTCCTGTCAGGAATGCCAGAACCTGGAAATC 435
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGlu 174
Db 616 AGAAAAACAGTCATAGACTGCAGCATCTCCAGTGAC---AAATTTGAATACCTTTTAAAT 672
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArgArg 249
Db 673 TTTGATAACACC---TTTGAGATCCACGACGACATAGAGTACTGAAGCGGATGGGAATG 729
QY 250 AsnSerIleGlyArgAlaGlyAlaThrLeu 260
Db 730 TCCTTTGGTCTGAGTCAGGCAAAATGCTCTCTG 762
```

RESULT 14

```
US-09-710-861-3
; Sequence 3, Application US/09710861
; Patent No. 6387649
; GENERAL INFORMATION:
; APPLICANT: La Thangue, Nicholas
; TITLE OF INVENTION: TRANSCRIPTION FACTOR DP-3 AND ISOFORMS THEREOF
; FILE REFERENCE: 620-54
; CURRENT APPLICATION NUMBER: US/09/710,861
; PRIOR APPLICATION NUMBER: US/09/189,627
; PRIOR FILING DATE: 1998-11-10
; PRIOR APPLICATION NUMBER: 08/723,415
; PRIOR FILING DATE: 1996-09-30
; PRIOR APPLICATION NUMBER: GB 9610195
; PRIOR FILING DATE: 1996-05-15
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 1154
```

```
; TYPE: DNA
; ORGANISM: mouse
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1107)
US-09-710-861-3
```

Alignment Scores:

```
Pred. No.: 1,78e-37 Length: 1154
Score: 384.00 Matches: 94
Percent Similarity: 57.37% Conservative: 50
Best Local Similarity: 37.45% Mismatches: 75
Query Match: 29.11% Indels: 32
DB: 3 Gaps: 9
```

US-10-088-830-2 (1-261) x US-09-710-861-3 (1-1154)

```
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 82 GTCACTCAGACTCAGACTGAGCTGCT-----111
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 112 -----GGCTGGTTCAGTAACGAGCAAGCAAAAGGAGATAA---150
QY 58 AlaAlaProArgIleThrGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 151 -----AATGGAAAGGCTTGAGACATTTTCAATGAAGGTGTGTGAGAAA 195
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 196 GTTCAGCGGAAGGCAACATTCATCAATGAGTAGCTGAGCTGGTATCTGAGTTT 255
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 256 ACCAACTCAAATAACCATCTGGCAGCTGATTTCGGCTTATGATCAGGAGAACATTAGACGA 315
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 316 AGAGTTTATGATGCTTTAAATGCTTAATGCGATGAACATAATTTCAAAGGAAAAA 375
QY 135 GluIleArgTyrMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGlu 154
Db 376 GAATCAAGTGGATGGCTGCTACCAATTCCTGTCAGGAATGCCAGAACCTGGAAATC 435
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGlu 174
Db 436 GAGAAAGCAGAGCGGATAGAACCGGATAAAGCAGAGCGAGCCAGCTACAAGAACTTCTC 495
QY 175 LysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAla 194
Db 496 CTTTCAGCAAAATGCTTTTAAAAAACCTGGTACAGAGAAATCGACAAAATGAACAAACAAAC 555
QY 195 GluAsn-----ValAsnGly---IleArgLeuProPheValLeuValLysThrSer 210
Db 556 CAGGGCCCTCCAGCTGTGAATTCACCATTCAGCTGCCATTTTATATCATTTAATACAGC 615
QY 211 ArgLysAlaArgValGluIleGluIleSerAspAspSerLysPheAlaHis---PheGlu 229
Db 616 AGAAAAACAGTCATAGACTGCAGCATCTCCAGTGAC---AAATTTGAATACCTTTTAAAT 672
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArgArg 249
Db 673 TTTGATAACACC---TTTGAGATCCACGACGACATAGAGTACTGAAGCGGATGGGAATG 729
QY 250 AsnSerIleGlyArgAlaGlyAlaThrLeu 260
Db 730 TCCTTTGGTCTGAGTCAGGCAAAATGCTCTCTG 762
```

RESULT 15

```
US-08-428-131-1
; Sequence 1, Application US/08428131
; Patent No. 5863757
```


GENERAL INFORMATION:
 APPLICANT: La Thangue, Nicholas Barrie
 TITLE OF INVENTION: Transcription Factor DP-1
 NUMBER OF SEQUENCES: 14
 CORRESPONDENCE ADDRESS:
 ADDRESSEE: Nixon & Vanderhye
 STREET: 1100 No. 5863757th Glebe Road, 8th Floor
 CITY: Arlington
 STATE: Virginia
 COUNTRY: U.S.A.
 ZIP: 22201-4714
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25 (EPO)
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/08/428,131
 FILING DATE: 23-JUN-1995
 CLASSIFICATION: 514
 ATTORNEY/AGENT INFORMATION:
 NAME: Arthur R. Crawford
 REGISTRATION NUMBER: 25,327
 REFERENCE/DOCKET NUMBER: 117-181
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (703) 816-4000
 TELEFAX: (703) 816-4100
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 1700 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 55..1284
 US-08-428-131-1

Alignment Scores:
 Pred. No.: 1.56e-36 Length: 1700
 Score: 378.50 Matches: 87
 Percent Similarity: 62.44% Conservative: 41
 Best Local Similarity: 42.44% Mismatches: 60
 Query Match: 28.70% Indels: 17
 DB: 2 Gaps: 7

US-10-088-830-2 (1-261) x US-08-428-131-1 (1-1700)

QY	50	LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu	69
DB	367	CGGACAGGAGGGCGGAGAG-----AATGGCAAGGGCTGGCGCAT	408
QY	70	TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal	89
DB	409	TTCTCCATGAAGGTGTGTGAGAGGTGCAGAGGAAAGGACCACTCTACAAATGAGGTG	468
QY	90	AlaAspGluIleTyrSerGluLeuLysSerMet---AlaHisIle-----GlyGln	105
DB	469	GCTGACGAGCTGTGGCAGAGTTTCAGCGCTGCAGCAACACATCTTACCAACGAATCA	528
QY	106	GlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAla	125
DB	529	GCTTATGACGAGAGAACATCCGGCGGCTGTCTACGATGCCTTAATGTGCTAATGGCC	588
QY	126	LeuArgValIleAlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArg	145
DB	589	ATGAACATCATCTCCAAGGAGAGAGGAGATCAATGGATCGCTGCCCACTCA	648
QY	146	TyrGluLysIleLysLysLeuGluValArgLysGluLeuValAsnLysIleArgAsn	165
DB	649	GCTCAGGAGTGCCAGAACTTAGAGGTGGAGGAGGAGGAGGAGGAGGATCAAAACAG	708

QY	166	LysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeu	185
DB	709	AGCAGTCTCAGCTCCAGGAGCTCATCTCGAGCAAAATTCCTTCAAGAACTTGGTGCAG	768
QY	186	ArgAsnGlnThrLeuGluSerSerAlaGluAsnVal-----AsnGlyIleArg	201
DB	769	AGAAATCGCAAGCTGAGCAGCAGGCGCGCCCTCTCCCAACTCTGTCAATCCAC	828
QY	202	LeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGluIleSerAsp	221
DB	829	TTGCCCTTCATCATTTGTCAACACACAGCAGCAGATCATTCATTCGACGATCTCCAAT	888
QY	222	AspSerLysPheAlaHis---PheGluPheAsnGlyAlaProPheThrLeuHisAspAsp	240
DB	889	GAC---AAATTTGAGTATCTGTTTAACTTTGACAAACAG---TTTGAGATCCACGATGAC	942
QY	241	LeuSerIleLeuGlu	245
DB	943	ATTGAGGTGCTCAAG	957

Search completed: February 7, 2005, 00:24:07
 Job time : 201 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 6, 2005, 23:22:07 ; Search time 580 Seconds
(without alignments)
2590.084 Million cell updates/sec

Title: US-10-088-830-2

Perfect score: 1319

Sequence: 1 MAPPRGAAATAALDLTG.....SILEGVRNRSIGRAGRATLH 261

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4313806 seqs, 2877871033 residues

Total number of hits satisfying chosen parameters: 8627612

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO.spool/US1008830/runat_04022005.122604_29641/app_query.fasta_1.455
-DB=Published Applications NA -QFMT=fastap -SUFFIX=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCLALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0
-MAXLEN=200000000 -USER=US1008830 @CGN 1 1 480 @runat_04022005.122604_29641
-NCPUS=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSFBLCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/2/pubpna/PCTU5_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/2/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/2/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/2/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq*
- 22: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Query Length	DB ID	Description
C 1	842	63.8	1406	18	US-10-437-963-33888
2	779.5	59.1	774	18	US-10-425-115-4115
3	492	37.3	1639	18	US-10-739-930-1190
4	464	35.2	1350	17	US-10-425-114-4049
5	464	35.2	1483	18	US-10-425-115-2033
6	463	35.1	1585	18	US-10-425-115-87351
7	450	32.6	1431	17	US-10-424-599-43806
8	399.5	30.3	2320	17	US-10-641-643-1456
9	394	29.9	1266	9	US-09-954-456-1176
10	394	29.9	2968	18	US-10-357-930-20983
11	394	29.9	2968	18	US-10-357-930-20983
12	387	29.3	2968	14	US-10-198-846-10136
13	379.5	28.8	476	18	US-10-767-701-17642
14	379.5	28.8	1751	17	US-10-424-599-43105
15	374.5	28.4	692	15	US-10-206-901B-21
16	374.5	28.4	1440	17	US-10-172-118-1357
17	374.5	28.4	1440	17	US-10-160-554-4
18	374.5	28.4	1440	17	US-10-342-887-1357
19	374.5	28.4	1440	17	US-10-641-643-1399
20	374.5	28.4	1440	18	US-10-717-597-73
21	374.5	28.4	2607	15	US-10-106-698-569
22	372.5	28.2	1213	18	US-10-437-963-63675
23	370	28.1	642	18	US-10-767-795-6106
24	361.5	27.4	2673	16	US-10-240-965-234
25	361	27.4	1167	17	US-10-425-114-20051
26	354	26.8	1832	9	US-09-220-091-8
27	342.5	25.0	911	17	US-10-425-114-36537
28	337.5	25.6	792	18	US-10-437-963-64593
29	299.5	22.7	1690	15	US-10-053-248-23
30	299.5	22.7	1890	18	US-10-345-837-23
31	261.5	19.8	358	18	US-10-856-499-566
32	254.5	19.3	364	18	US-10-856-499-465
C 33	252.5	19.1	505	16	US-10-029-386-8935
C 34	227.5	17.2	690	11	US-09-969-034-3453
C 35	227.5	17.1	738	17	US-10-424-599-91931
36	213	16.1	774	18	US-10-425-115-88311
C 37	188.5	14.3	425	9	US-09-815-343-1294
C 38	188.5	14.3	425	17	US-10-097-105-1294
39	187.5	14.2	352	18	US-10-856-499-440
40	185.5	14.1	432	9	US-09-815-343-12
41	185.5	14.1	432	17	US-10-097-105-12
C 42	180	13.6	411	9	US-09-998-598-1532
43	167.5	12.7	488	17	US-10-242-535A-27279
44	167.5	12.7	488	17	US-10-085-783A-27279
45	146.5	11.1	271	18	US-10-856-499-576

ALIGNMENTS

RESULT 1

US-10-437-963-33888/c
; Sequence 33888, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(5321)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 33888

```

Qy 170 LeuGlnGluIleGlyLysGlnPheAspLeuGlnAseNileLysLeuArgAsnGlnThr 189
Db 575 CTCAGGAAATTTGAAAACGCGTTTGATCACCTTCAGATAATTACATTACGCCAACCCAGGCT 516
Qy 190 LeuGluSerSerAlaGluAsnValAsnClylleArgLeuProPheValLeuVallysThr 209
Db 515 AGTCAAGAGCCAGCAGAGAAGTGTAAATGGCATCTCTTCCTTCGGTCTTATTGATCAAGACA 456
Qy 210 SerArgLysAlaAqValGluIleGluIleSerAspSerLysPheAlaHisPheGlu 229
Db 455 TCCGMAAAGCAAGGTGGAATTCAGATTTCCGAGATTCCAAGTTTGCACGGTTCGAC 396
Qy 230 PheAsnGlyAlaProPheThrLeuHisAspLeuSerilleLeuGluGlyValArgArg 249
Db 395 TTCAACGGTGCACCATTCAACATGCATGATGATCAATCCTTGAAGCCATCAGGCGT 336
Qy 250 AsnSerilleGlyArgAlaGly 256
Db 335 AACCAAAAGGAAGCGTGGC 315

RESULT 2
US-10-425-115-4115
; Sequence 4115, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 4115
; LENGTH: 774
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_103754C.1
US-10-425-115-4115

Alignment Scores:
Pred. No.: 1.58e-72 Length: 774
Score: 779.50 Matches: 166
Percent Similarity: 77.73% Conservative: 37
Best Local Similarity: 64.84% Mismatches: 43
Query Match: 59.10% Indels: 10
DB: 18 Gaps: 4

US-10-088-830-2 (1-261) x US-10-425-115-4115 (1-774)

Qy 4 ProArgGlyGlyAlaAlaAlaAlaThraAlaLeuAspLeuThrGlyValHisIle 23
Db 6 CCTTCGCGCATGCTCGCGGCGAGCC-----GAGCTCACCGCGCTCGGCATC 53
Qy 24 LeuGluAlaSerSerValProLeuPro-----GluAlaGlyGlyAsnAlaVal 40
Db 54 AGCGAAGGAGCTGGGATCCCTTCGCTTCCCGAGCGCGAAGCGTCGGCGCAATGCAGCT 113
Qy 41 GlnArgLysGlyAlaValAspProAspLysAspArgLysGlyLysAlaAlaPro 60
Db 114 GAGGGAAGAGGAGGGGTGTAAAGAACAAAGAGCGAGGAGGTGAAGGCTGGTAGCGAG 173
Qy 61 ArgIleThrGlyTrpGlyLeuArgGlyTyrrSerLysIleValCysGluLysValGluAla 80
Db 174 CGGATCCCGGGTGGGGGCTCCGCGAGTTCAGCAAGATAGTTTTCTAAAAAGTTGAGACC 233
Qy 81 LysGlyArgThrThrTyrrAsnGluValAlaAspGluIleTyrrSerGluLeuLysSerMet 100
Db 234 AAAGCAAGGACTACTTACAGTGAGTTGCGATCAAAATTTGTGAGAGCTGAAG---TTG 290

```


APPLICANT: Kovalic, David K.
APPLICANT: Zhou, Yihua
APPLICANT: Cao, Yongwei
TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
FILE OF INVENTION: Plants
FILE REFERENCE: 38-21(53222)B
CURRENT APPLICATION NUMBER: US/10/425,115
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 369326
SEQ ID NO 87351
LENGTH: 1585
TYPE: DNA
ORGANISM: Zea mays
FEATURE:
OTHER INFORMATION: Clone ID: MRT4577_179669C.1
US-10-425-115-87351

Alignment Scores:
Pred. No.: 9.39e-39 Length: 1585
Score: 463.00 Matches: 112
Percent Similarity: 54.80% Conservative: 42
Best Local Similarity: 39.86% Mismatches: 79
Query Match: 35.10% Indels: 48
DB: 18 Gaps: 8

US-10-088-830-2 (1-261) x US-10-425-115-87351 (1-1585)

QY 5 ArgGlyGlyAlaAlaAlaAlaAlaAlaAlaAlaAlaLeuThrGlyValHisLeu 24
DB 4 CGAGGCGAGCAGCGTCTCCGCTTGAATATCTGCATCAACGCGC----- 51
QY 25 GluAlaSerSerValProLeuProGluAlaGlyAsnAlaValGlnArgLysGly 44
DB 52 -----GACGACCGCGCTCGTCCGAGCTCTACGAGCAAGAAAGAGAGGC 102
QY 45 -----AlaValAspProAspLysAspArgLysGlyGluLysAlaAlaProArgIle 62
DB 103 ACACGGCGAGTGGTCTCTAATAAGGTAACCGG----- 135
QY 63 ThrGlyTrpGlyLeuArgGluTrpSerLysIleValCysGluLysValGluAlaLysGly 82
DB 136 -----GGACTGCGCCAGTTTATGATGAAAGTTGTGAGAAAGTTGAAAGTAAAGG 186
QY 83 ArgThrThrThrAsnGluValAlaAspGluIleTrpSerGluLeuLysSerMetAlaHis 102
DB 187 AGACAAACATATATGAGTGGCGAGTGAATCTGTGCTGAGTTTACAGACCCCAACAT 246
QY 103 -----ileGlyGlnGlyPheAspGluLysAsnIle 112
DB 247 AATATTGAGGCACAGATCTGTATAACCTTAACGCGCAACAATATGATGAGAAATAATA 306
QY 113 ArgArgArgValTrpAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGlu 132
DB 307 CGACGAAGAGTTATGATCTTTGAATGTGTGATGCTATGACATTTATCTAAAGAT 366
QY 133 LysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeu 152
DB 367 AAAAGGAGATCCAGTGGAGGCTTGGCGGT-----ACTAGTATAGTGACATT 417
QY 153 GluGluValArgLysGluLeuVal-----AsnLysIleArgAsnLysLysAlaLeu 169
DB 418 GAAGAATGAAGACGCGAGCTTGTGGGACTGAAGAGTAGGATTGAAAGAAAGTAGCTTAC 477
QY 170 LeuGlnGluIleGluLysGlnPheAspLeuGlnAsnIleLysLeuArgAsn---Gln 188
DB 478 CTACAGGAGCTACAAAGCAATATGTAGTCTGCAAAACCTGATTCAACGAATGAGCAA 537
QY 189 ThrLeuGluSerSerAlaGluAsnValAsnGlyIleArgLeuProPheValLeuValLys 208
DB 538 TCATATGTTACGAAACAGCGCTTCTGGTGGAGTGGCTTGGCAATTTATCTAGTTAC 597
QY 209 ThrSerArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHisPhe 228

DB 598 ACCGACCTCATGCTACCGTGGAAAGTTGAGATATCAGAGATATCAGAGCTGGTGCATTTT 657
QY 229 GluPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArg 248
DB 658 GACTTCAATAGCACCCCAATTCGAGCTGCAGCAGCACTCATACGCTCTTAAAGAAATCGGA 717
QY 249 -----ArgAsnSerIleGlyArgAlaGlyArgAla 258
DB 718 TTCTGTGGAGAGAACACATGACAGCACTCAAGAGTCGATATCAATGGAGGTGAGAGC 777
QY 259 Thr 259
DB 778 TCA 780

RESULT 7
US-10-424-599-43806
Sequence 43806, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 43806
LENGTH: 1431
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT MRT3847_139556C.1
US-10-424-599-43806

Alignment Scores:
Pred. No.: 2.45e-35 Length: 1431
Score: 430.00 Matches: 105
Percent Similarity: 59.84% Conservative: 41
Best Local Similarity: 43.03% Mismatches: 73
Query Match: 32.60% Indels: 26
DB: 17 Gaps: 4

US-10-088-830-2 (1-261) x US-10-424-599-43806 (1-1431)

QY 27 SerSerValProPro----- 31
DB 155 AACAGCTTCCACCTCCCGTCGATAAACTATTCTCAAAATTGAACCATCTCGAGTACAT 214
QY 32 LeuProGluAlaGlyCysAsnAlaValGlnArgLysGlyAlaValAspProAspLysAsp 51
DB 215 GCTGATGATGCTGGATGCACACGCTTCTCTTCCAGTGGC-----AAGAAAGAAAG 265
QY 52 ArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSer 71
DB 266 AAGGCGGTCAACGGCTTGGACCTGATAGAGTGGAGAGGTCTCCGCCAATTAGT 325
QY 72 LysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAsp 91
DB 326 ATGAAAGTGTGTGAGAGGTAGAAAGCAGCGGGGAAGAACACACATACATAGAGTGGCGAT 385
QY 92 GluIleTyrSerGluLeuLysSerMetAlaHisIle-----GlyGlnGly 106
DB 386 GAGCTTGTGCTGAAATTTCTGAACCAAGCAATAGTAATGCGCCCTGATCAGCAACAA 445
QY 107 PheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAlaLeu 126
DB 446 TATGATGAAAAAACAATCCCGGAGGTCTATGATGCTCTGACGTTTCTCATGGCAATG 505
QY 127 ArgValIleAlaLysGluLysLysGluIleArgTrpMetGlyLeuSerAsnTyrArgTyr 146

Db 506 GATATTATTTCCAGGACAAAAGAAATTCATGGAGGGCCCTCTCGTACTACTGTG 565
Qy 147 GlulyslleLysLysLeuGluValargLysGluValAsnLyslleArgAsnLys 166
Db 566 AATGATATTGAAGAACTAAAGACAGAGCGGCTTGGGCTCAGGAATAGAAATGAAAGAA 625
Qy 167 LysAlaLeuGluGluGluLysGluPheAspLeuGluAsnLyslleLysLeuAr 186
Db 626 ACAGCCTATCTGCAGGAGCTTGAGGGAGCAATTCATAGGCTTCAGAACCTTATTCAACG 685
Qy 186 gAsnGlnThrLeuGluSerSerAlaGluAsnValAsnGly---lleArgLeuProPheVa 205
Db 686 AAACGAGCAGTTATAGCTCAAGAAATCTCCAGTGGAGGTATCTTACCTTTAT 745
Qy 205 lleuValLysThrSerArgLysAlaArgValGluLyslleGluLysSerAspSerLysPh 225
Db 746 -TTGGTACAGACAGCCCTCATGCAACTGGGGGAATGGAATATCAGAGATATGCAGCT 804
Qy 225 eAlaHiePheGluPheAsnGlyAlaPropheThrLeuHisAspLeuSerlleLeuG1 245
Db 805 TGTTCATTTTGAATTCATAGCACTCTTTTGAGCTGCATGACGACAAATATGTTCTCAA 864
Qy 245 uGlyValarg 248
Db 865 GGCAATGAAA 874

RESULT 8

US-10-641-643-1456
; Sequence 1456, Application US/10641643
; Publication No. US20040077003A1
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; Susan G. Stuart
; Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL
; GENE EXPRESSION

NUMBER OF SEQUENCES: 1508
CORRESPONDENCE ADDRESS:
ADDRESS: INCYTE PHARMACEUTICALS, INC.
STREET: 3174 PORTER DRIVE
CITY: PALO ALTO
STATE: CALIFORNIA
COUNTRY: USA
ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/641.643
FILING DATE: 14-Aug-2003
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: <Unknown>
FILING DATE: <Unknown>
ATTORNEY/AGENT INFORMATION:
NAME: Zeller, Karen J.
REGISTRATION NUMBER: 37,071
REFERENCE/DOCKET NUMBER: PA-0001 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (650) 855-0555
TELEFAX: (650) 845-4166
INFORMATION FOR SEQ ID NO: 1456:
SEQUENCE CHARACTERISTICS:
LENGTH: 2320 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: GENBANK
CLONE: 9604478

SEQUENCE DESCRIPTION: SEQ ID NO: 1456 :

US-10-641-643-1456

Alignment Scores:
Pred. No.: 7,74e-32 Length: 2320
Score: 399.50 Matches: 94
Percent Similarity: 58.33% Conservative: 53
Best Local Similarity: 37.30% Mismatches: 88
Query Match: 30.29% Indels: 17
DB: 17 Gaps: 7

US-10-088-830-2 (1-261) x US-10-641-643-1456 (1-2320)

Qy 18 LeuThrGlyValHisLysLeuGluAlaSerSerValProProLeuProGluAlaGlyGly 37
Db 222 GTTACTCAGACACACATAGCAGAACTACTGGTGGTGCCT-----GGT 266
Qy 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 267 GATAGAAAACGGGCTAGAAAATTTATAGACTCTGATTTTTCAGAAAGTAAACGAGCAA 326
Qy 58 AlaAlaProArgLysThrGlyTrpGlyLeuArgGluLysSerLysLysValCysGluLys 77
Db 327 AAAGGAGATAAA---AATGGAAAGGCTTGAGACACTTTTCAATGAAGTGTGTGAGAA 393
Qy 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluLysTyrSerGluLeu 97
Db 384 GTTCAACGAAAAGGTACAACTCGTACAATGAATGCTGATGCTGATGCTGATGCTGATGCT 443
Qy 98 LysSer-----MetAlaHisLysGlyGlnGlyPheAspGluLysAsnLysAla 113
Db 444 ACCAATTCAAATAACCAATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 503
Qy 114 ArgArgValTyrAspAlaPheAsnValLeuLysAlaLeuArgValLysAlaLysGluLys 133
Db 504 CGAAGAGTTTATGATGCTTTAAATGTCTAATGTCGCAATGAACATAATTTCAAAGGAAAA 563
Qy 134 LysGluLysArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLysLysLys 153
Db 564 AAAGAAATCAAGTGGATTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 623
Qy 154 GluValArgLysGluLeuValAsnLysLysLysLysLysLysLysLysLysLysLysLysLys 173
Db 624 ATAGAGACGACGAGCGGATAGAACGGATAAGCAGACAGCGGCGCCAGCTGCAAGACTT 583
Qy 174 GluLysGlnPheAspLeuGlnAsnLysLysLysLysLysLysLysLysLysLysLysLysLys 193
Db 684 CTCCTACAGCAAAATCGCTTTCAAAAACCTGGTACAGAGAAATCGACAAATGACGACGAA 743
Qy 194 AlaGluAsnVal-----AsnGlyLysArgLeuProPheValLeuValLysThr 209
Db 744 AACGAGGCGCGCGCTGCTGAACTCTACCATTCAGCTGCCATTCATATCATCATCAATACA 803
Qy 210 SerArgLysAlaArgValGluLysLysLysLysLysLysLysLysLysLysLysLysLysLys 228
Db 804 ASCAGAAAAACAGTCATAGATTGACAGCTCTCCAGTGAC---AAGTTTTCAGTATCTTTTC 860
Qy 229 GluPheAsnGlyAlaPropheThrLeuHisAspLeuSerLysLysLysLysLysLysLys 248
Db 861 AATTTTGACACACC---TTTGAGATCCATGATGATGATGATGATGATGATGATGATGATGAT 917
Qy 249 ArgAsnSerLysGlyArgAlaGlyArgAlaThrLeu 260
Db 918 ATGTCGTTGGCTGGAGTCAGGCAAAATGCTCTCTG 953

RESULT 9

US-09-954-456-1176
; Sequence 1176, Application US/09954456
; Patent No. US20020115057A1
; GENERAL INFORMATION:
; APPLICANT: Young, Paul
; TITLE OF INVENTION: Process for Identifying Anti-Cancer Therapeutic Agents Using Canc
; TITLE OF INVENTION: Sets
; FILE REFERENCE: 689290-76


```
; CURRENT APPLICATION NUMBER: US/09/954,456
; CURRENT FILING DATE: 2001-09-18
; PRIOR APPLICATION NUMBER: US/60/233,617
; PRIOR FILING DATE: 2000-09-18
; PRIOR APPLICATION NUMBER: US/60/234,052
; PRIOR FILING DATE: 2000-09-20
; PRIOR APPLICATION NUMBER: US/60/234,923
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,134
; PRIOR FILING DATE: 2000-09-25
; PRIOR APPLICATION NUMBER: US/60/235,637
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,638
; PRIOR FILING DATE: 2000-09-26
; PRIOR APPLICATION NUMBER: US/60/235,711
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,720
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,840
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: US/60/235,863
; PRIOR FILING DATE: 2000-09-27
; NUMBER OF SEQ ID NOS: 2276
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 1176
; LENGTH: 1266
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-954-456-1176

Alignment Scores:
Pred. No.: 1,296-31 Length: 1266
Score: 394.00 Matches: 94
Percent Similarity: 58.17% Conservatives: 52
Best Local Similarity: 37.45% Mismatches: 89
Query Match: 29.87% Indels: 16
DB: Gaps: 8

US-10-088-830-2 (1-261) x US-09-954-456-1176 (1-1266)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37
Db 114 GTTACTCAGACACACATAGCAGAACTACTGGTGGTCCCT-----GGT 158
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 159 GATAGAAAACGGCTAGCAAAATTTATAGACTCTGATTTTCAGAAAGTAAACGAGCAA 218
QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTyrSerLysIleValCysGluLys 77
Db 219 AAAGGAGATAAA---AATGGGAAGGCTTGACACACTTTTCAATGAAAGTGTGAGAAA 275
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluIleTyrSerGluLeu 97
Db 276 GTTCAACGAAAAGGTACAAACATCGTACATGAATGAATCGCTGATGAGCTGGTGCAGAGTTC 335
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 336 ACCAATTCAAATAACCATTTGGCTGCTGATTCGGCTTATGATCAGAAAGAACATTTAGCGCA 395
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 396 AGAGTTTATGATGCTTTAATATGTCTTAATGGCAATGAACATATTTCAAGGAGAAAAA 455
QY 135 GluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysIleLysLysLeuGluGlu 154
Db 456 GAAATCAAGTGGATTCCTGCTACCAATTTCTGCTCAGGAATGTTCAGAACTCTGGAGATA 515
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGlu 174
Db 516 GAGAACGACAGCGGATGAGCAAGCAAGTAAAGCAGCGGCCCGCTGCAAGAACTTC 575
QY 175 LysGlnPheAspAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerAla 194
```

```
Db 576 CTACAGCAAAATCGCTTTTCAAAAACCTGGTACAGAGAAATCGACAAATGAGCAGCAAAAC 635
QY 195 GluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSer 210
Db 636 CAGGCGCCCGCGCTCTGAACCTTACCATTCAGCTGCCATTCATAATCATCAATCAACAGC 695
QY 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheGlu 229
Db 696 AGAAAAACAGTCATAGATTGCGAGCATCTCCAGTGAC---AAGTTTGAGTATCTTTTCAAT 752
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeuGluGlyValArgArg 249
Db 753 TTTGACAAACACC---TTTGAGATCCATGATGACATAGAAAGTACTAAAGCGGATGGAATG 809
QY 250 AsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 810 TCGTTTGGCTGGAGTCAGGCAAAATGCTCTCTG 842

RESULT 10
US-10-357-930-20983
; Sequence 20983, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlögel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; TITLE OF INVENTION: IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; TITLE OF INVENTION: HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 20983
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-20983

Alignment Scores:
Pred. No.: 4,1e-31 Length: 2968
Score: 394.00 Matches: 94
Percent Similarity: 58.17% Conservatives: 52
Best Local Similarity: 37.45% Mismatches: 89
Query Match: 29.87% Indels: 16
DB: Gaps: 8

US-10-088-830-2 (1-261) x US-10-357-930-20983 (1-2968)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerValProLeuProGluAlaGlyGly 37
Db 256 GTTACTCAGACACACATAGCAGAACTACTGGTGGTCCCT-----GGT 300
```

```
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 301 GATAGAAAACGGCTAGAAATTTATAGACTCTGATTTTCAGAAAGTAAACGAGCAAA 360
QY 58 AlaAlaProArgLysGlyTrpGlyLeuArgGlnLysSerLysLysValCysGluLys 77
Db 361 AAGGAGATAAAA---AATGGAAAGGCTTGAGACACTTTTCAATGAAAGTGTGTGAGAAA 417
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluLysLysSerGluLeu 97
Db 418 GTTCAACGAAAGGTACAACATCGTACAATGAATCGCTGATGAGCTGGTGTGAGAGTTC 477
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 478 ACCAATTCAAATACCAATTTGGCTGCTGATTCGGCTTATGATCAGAGAAACATTAGCGGA 537
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 538 AGAGTTTATGATGCTTTAAATGTCTAATGGCAATGAACATAATTTCAAGGAAAGAAAAA 597
QY 135 GluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLysLysGlu 154
Db 598 GAAATCAAGTGGATGGCTGCTTACCAATTTCTGCTCAGGAATGTCAGAAATCTGGAGATA 657
QY 155 ValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuGlnGluIleGlu 174
Db 658 GAGAGCAGAGCGGCTAGAACGATGAATGAACAGAGCGGCGCCAGCTGCAAGAACTTCTC 717
QY 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAla 194
Db 718 CTACAGCAAAATCGCTTTCAAAAACCTGGTACAGAGAAATCGACAAATCGAGCAAGCA 777
QY 195 GluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSer 210
Db 778 CAGGCGCCCGGCTCTGAACTCTACCATTCAGCTGCCATTATATATCATATCATACAGC 837
QY 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheGlu 229
Db 838 AGAAAACACATGATAGTTCAGCACTCCAGTGAC---AAGTTTGAGTATCTTTTCAAT 894
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLysSerIleLeuGluGlyValArg 249
Db 895 TTTGACAACACC---TTTGAGATCCATGATGACATAGAGTACTAAAGCGGATGGGAATG 951
QY 250 AsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
Db 952 TCGTTTGGCTGGAGTCAGGCAAAATGCTCTCTG 984

RESULT 11
US-10-357-930-26827
; Sequence 26827, Application US/10357930
; Publication No. US20040259086A1
; GENERAL INFORMATION:
; APPLICANT: Schlegel, Robert
; APPLICANT: Endege, Wilson
; APPLICANT: Monahan, John
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
; IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY OF
; HUMAN PROSTATE CANCER
; FILE REFERENCE: MRI-007BCN
; CURRENT APPLICATION NUMBER: US/10/357,930
; CURRENT FILING DATE: 2003-02-04
; PRIOR APPLICATION NUMBER: 09/785,276
; PRIOR FILING DATE: 2003-02-16
; PRIOR APPLICATION NUMBER: 60/183,319
; PRIOR FILING DATE: 2000-02-17
; PRIOR APPLICATION NUMBER: 60/189,862
; PRIOR FILING DATE: 2000-03-16
; PRIOR APPLICATION NUMBER: 60/207,454
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: 60/211,314
; PRIOR FILING DATE: 2000-06-09
; PRIOR APPLICATION NUMBER: 60/219,007
```

```
; PRIOR FILING DATE: 2000-07-18
; PRIOR APPLICATION NUMBER: 60/255,281
; PRIOR FILING DATE: 2000-12-13
; NUMBER OF SEQ ID NOS: 62232
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26827
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 1..2
; OTHER INFORMATION: n = A,T,C or G
US-10-357-930-26827

Alignment Scores:
Pred. No.: 4,1e-31 Length: 2968
Score: 394.00 Matches: 94
Percent Similarity: 58.17% Conservative: 52
Best Local Similarity: 37.45% Mismatches: 89
Query Match: 29.87% Indels: 16
DB: 18 Gaps: 8

US-10-088-830-2 (1-261) x US-10-357-930-26827 (1-2968)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
Db 256 GTTACTCAGACACACATAGACAGACTCTGCTGGTCCCT-----GGT 300
QY 38 AsnAlaValGlnArgLysGlyAlaValAspProAspLysAspArgLysLysGluLysAla 57
Db 301 GATAGAAAACGGCTAGAAATTTATAGACTCTGATTTTCAGAAAGTAAACGAGCAAA 360
QY 58 AlaAlaProArgLysGlyTrpGlyLeuArgGluTyrSerLysLysValCysGluLys 77
Db 361 AAGGAGATAAAA---AATGGAAAGGCTTGAGACACTTTTCAATGAAAGTGTGTGAGAAA 417
QY 78 ValGluAlaLysGlyArgThrThrTyrAsnGluValAlaAspGluLysLysSerGluLeu 97
Db 418 GTTCAACGAAAGGTACAACATCGTACAATGAATGCTGATGAGCTGGTGTGAGAGTTC 477
QY 98 ---LysSerMetAlaHisIleGly-----GlnGlyPheAspGluLysAsnIleArgArg 114
Db 478 ACCAATTCAAATACCAATTTGGCTGCTGATTCGGCTTATGATCAGAGAAACATTAGCGGA 537
QY 115 ArgValTyrAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
Db 538 AGAGTTTATGATGCTTTAAATGTCTAATGGCAATGAACATAATTTCAAGGAAAGAAAAA 597
QY 135 GluIleArgTrpMetGlyLeuSerAsnTyrArgTyrGluLysLysLysLysLysGlu 154
Db 598 GAAATCAAGTGGATGGCTGCTTACCAATTTCTGCTCAGGAATGTCAGAAATCGAGCAAG 657
QY 155 ValArgLysGluLeuValAsnLysLysLysLysLysLysLysLysLysLysGlu 174
Db 658 GAGAGCAGAGCGGCTAGAACGATGAATGAACAGAGCGGCGCCAGCTGCAAGAACTTCTC 717
QY 175 LysGlnPheAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAla 194
Db 718 CTACAGCAAAATCGCTTTCAAAAACCTGGTACAGAGAAATCGACAAATCGAGCAAGCA 777
QY 195 GluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSer 210
Db 778 CAGGCGCCCGGCTCTGAACTCTACCATTCAGCTGCCATTATATATCATATCATACAGC 837
QY 211 ArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheGlu 229
Db 838 AGAAAACACATGATAGTTCAGCACTCCAGTGAC---AAGTTTGAGTATCTTTTCAAT 894
QY 230 PheAsnGlyAlaProPheThrLeuHisAspAspLysSerIleLeuGluGlyValArg 249
Db 895 TTTGACAACACC---TTTGAGATCCATGATGACATAGAGTACTAAAGCGGATGGGAATG 951
```

```
QY 250 AasSerIleGlyArgAlaGlyArgAlaThrLeu 260
   ::::
Db 952 TCGTTGGCTGGAGTCAGGCAAAATGCTCTCTG 984

RESULT 12
US-10-198-846-10136
; Sequence 10136, Application US/10198846
; Publication No. US2003099974A1
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10136
; LENGTH: 2968
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 1, 2
; OTHER INFORMATION: n = A,T,C or G
US-10-198-846-10136

Alignment Scores:
Pred. No.: 2,248-30 Length: 2968
Score: 387.00 Matches: 95
Percent Similarity: 58.33% Conservative: 52
Best Local Similarity: 37.70% Mismatches: 88
Query Match: 29.34% Indels: 17
DB: 14 Gaps: 8

US-10-088-830-2 (1-261) x US-10-198-846-10136 (1-2968)
QY 18 LeuThrGlyValHisIleLeuGluAlaSerSerValProLeuProGluAlaGlyGly 37
   ::::
Db 256 GTTACTCAGACACATACAGAGCTACTGGTGGTCCCT-----GGT 300

QY 38 AenAlaValGlnArgLysGlyAlaValAaspProAaspLysArgLysLysGluLysAla 57
   ::::
Db 301 GATAGAAAACGGCGTAGAAAATTTATAGACTCTGATTTTTCAGAAAAGTAACGAAGCAAA 360

QY 58 AlaAlaProArgIleThrGlyTrpGlyLeuArgGluTySerLysIleValCysGluLys 77
   ::::
Db 361 AAAGGAGATAAA---AATCGGAAGGCTTGAGACACTTTTCAATGAAGAGTGTGAGAAA 417

QY 78 ValGluAlaLysGlyArgThrThrTyAsnGluValAlaAaspGluIleTySerGluLeu 97
   ::::
Db 418 GTTCAACGAAAGGTACACATCGTACATGAATGAATGCTGATGAGTGGTGTGCAGATTTC 477

QY 98 ---LysSerMetAlaHisIleGly-----GlnGly-PheAaspGluLysAsnIleArg 114
   ::::
Db 478 ACCAATTCAATAAACCATTTGGTGTCTGATTCGGGCTTATGATCAGAAACATTAGGGC 537

QY 114 GargValTyAspAlaPheAsnValLeuIleAlaLeuArgValIleAlaLysGluLysLys 134
   ::::
Db 538 AAGAGTTTATGATGCTTTAAATGTGCTAATGCAATGAACATAATTTCAAGGAAAAAAA 597

QY 134 sgluileArgTtpMetGlyLeuSerAsnTyArgTyGluLysIleLysLysLeuGluGl 154
   ::::
Db 598 AGAATCAAGTGGATTGGCTGCCTACCAATTCCTCAGGATGTCAGAAATCTGGAGAT 657

QY 154 uValArgLysGluLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnIleGl 174
   ::::
   ::::
```

```
Db 658 AGAGAAGCAGAGCGCGGATAGAACGCGATAAAGCAGAAAGCGGCGCCAGTCGCAAGAACTTCT 717
QY 174 ulysGlnPheAaspAspLeuGluAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAl 194
   ::::
Db 718 CCTACAGCAAAATCGCTTTCAAAACCTGGTACAGAAATCGCAAAAATGAGCAGCAAAA 777

QY 194 aGluAsnVal-----AsnGlyIleArgLeuProPheValLeuValLysThrSe 210
   ::::
Db 778 CCAGGCGCGCGCGCTCTGAACCTCTACCATTCAGCTGCCATTCAATCATCAATCAACAG 837

QY 210 rArgLysAlaArgValGluIleGluIleSerAspSerLysPheAlaHis---PheGl 229
   ::::
Db 838 CAGAAAACAGTCATAGATTCGAGCATCTCCAGTGAC---AAGTTTGAGTATCTTTTCAA 894

QY 229 uPheAsnGlyAlaProPheThrLeuHisAaspLeuSerIleLeuGluGlyValArgAr 249
   ::::
Db 895 TTTTGACAACAC---TTTGAGATCCTCATGATAGTAGTACTAAAGCGGATGGGAAT 951

QY 249 gAsnSerIleGlyArgAlaGlyArgAlaThrLeu 260
   ::::
Db 952 GTCGTTTGGCTGGAGTCAGGCAAAATGCTCTCTG 985

RESULT 13
US-10-767-701-17642
; Sequence 17642, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 17642
; LENGTH: 476
; TYPE: DNA
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3476-039-P1-K1-D2
US-10-767-701-17642

Alignment Scores:
Pred. No.: 1,168-30 Length: 476
Score: 379.50 Matches: 79
Percent Similarity: 84.21% Conservative: 17
Best Local Similarity: 69.30% Mismatches: 16
Query Match: 28.77% Indels: 2
DB: 18 Gaps: 1

US-10-088-830-2 (1-261) x US-10-767-701-17642 (1-476)
QY 139 MetGlyLeuSer-AsnTyArgTyArgLysIleLysLysLeuGluGluValArgLysGl 158
   ::::
Db 3 ATGGGCTTGTCTACACTTCCATATGAATACATA---AAGTTGGAGAAATCTGCCAAGA 59

QY 158 uLeuValAsnLysIleArgAsnLysLysAlaLeuLeuGlnGluIleGluLysGlnPheAs 178
   ::::
Db 60 ACTCATGATTAGGTTTGAACAAGAAAAAACTTCTCCAGGAAATTTGAACGACAGTTTGA 119

QY 178 pAspLeuGlnAsnIleLysLeuArgAsnGlnThrLeuGluSerSerAlaGluAsnValAs 198
   ::::
Db 120 TCACCTCCAGAAATCAAGTTTTCGCAACAGTTTACTACAGAGCGCCAGCTGAGAGTCGGA 179

QY 198 nGlyIleArgLeuProPheValLeuValLysThrSerArgLysAlaArgValGluIleGl 218
   ::::
Db 180 TGGTATCGCTTCTCCATCTTATTGGTCAAGCATCCAGAAAAGCAAGGGTGAATTTGA 239

QY 218 ulsSerAaspSerLysPheAlaHisPheGluPheAsnGlyAlaProPheThrLeuHi 238
   ::::
Db 240 GATTTTCAGAGAACTCAAGTTTGGTTCGACTTCAACTGCTACACCATTCACCTTGCA 299
```

```
QY 238 sAspLeuSerIleLeuGluGlyValArgArgAsnSer 251
Db 300 TGACGATGTCCTCAATCTTGAAGCGATCAGGTGTAATAAT 339

RESULT 14
US-10-424-599-43105
; Sequence 43105, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 43105
; LENGTH: 1751
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_138923C.1
US-10-424-599-43105

Alignment Scores:
Pred. No.: 6,76e-30 Length: 1751
Score: 379.50 Matches: 96
Percent Similarity: 47.89% Conservative: 40
Best Local Similarity: 33.80% Mismatches: 63
Query Match: 28.77% Indels: 86
DB: 17 Gaps: 3

US-10-088-830-2 (1-261) x US-10-424-599-43105 (1-1751)

QY 50 LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu 69
Db 299 AAAAGAGCGGACACACCGGCTGGTGACCTGATAAGAGTGAAGATGTCCTCCGCTAT 358
QY 70 TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal 89
Db 359 TTTAGTAGAAGTGTGTGAGAGGTAGAAAGCAGGGAAGGAGATCATACATGAGGTG 418
QY 90 AlaAspGluIleTyrSerGlu----- 96
Db 419 GCAGCTGAGCTTGTGCGAGAATGTAATAATAATGTTTCCACAAGTTCTATTGTATGGA 478
QY 96 ----- 96
Db 479 AATGTTTCTGAAAACAGAAACATCCAACTGAGAGAAAAAGAAACTGGTGAGAAATAAC 538
QY 96 ----- 96
Db 539 GTCTTTCTTTGTAGTTCAAAATTCACAAAAATTCGTGGTTGGATTCTCAATTTAGTA 598
QY 96 ----- 96
Db 599 CATATACCTAGTTATTTTAAAAAATTTCAATATGGATTGTGGTTATATTTAGGTAGCTG 658
QY 97 -----LeuLysSerMetAlaHis-----IleGlyGln 105
Db 659 ATGAACCTGTGGCGAATTTGTGTATCCCAATCAATGGTTTCAACTCTCGATCAG-CAA 717
QY 106 GlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAla 125
Db 718 CAATATGATCAGAAAAACATTCCTGCAAGGGTGTATGATGCTCTGAATGTTCTCATGGCA 777
QY 126 LeuArgValIleAlaLysGluLysGluIleArgTyrMetGlyLeuSerAsnTyrArg 145
Db 778 ATGATATATTATTTCTAAGGAAAAAGGAAATTCAAATGGAAGGGTCTCCCTCGTACTAGT 837

QY 146 TyrGluLysIleLysLysLysLeuGluValArgLysGluLeuValAsnLysIleArgAsn 165
Db 838 CTAAGCGATATTGAAGAGCTAAAGTCAGAGCGTCTTTGGGCTCAGGAATAGAAATGAAAG 897
QY 166 LysLysAlaLeuLeuGlnGluIleGluLysGlnPheAspAspLeuGlnAsnIleLysLeu 185
Db 898 AAAGCAGCCTATTGTCAAGAGCTGGAGGAGCAATACATAGGTCTTCAGAACTTATTCAA 957
QY 186 ArgAsnGlnThrLeuGluSerSerAlaGluAsnValAsn---GlyIleArgLeuProPhe 204
Db 958 CGAAATGAGCAACTATATAGCTCAGGAATGCTCCCAATGGAGGTGTATCTTTGCCCTTT 1017
QY 205 ValLeuValLysThrSerArgLysAlaA-gValGluIleLysSerAspSerLys 224
Db 1018 ATCTGTGTGAGACACGCGCCCATGCAACTGTTGAAGTGAATAATCAGAAATATGCAG 1077
QY 225 PheAlaHisPheLeuPheAsnGlyAlaProPheThrLeuHisAspAspLeuSerIleLeu 244
Db 1078 CTTGTGCATTTTGAATATATTTCGCTCCCTTTGAGATGCATGATGATAACTATGTCTC 1137
QY 245 GluGlyValArg 248
Db 1138 AAGCAATGAAA 1149

RESULT 15
US-10-206-901B-21
; Sequence 21, Application US/10206901B
; Publication No. US20030100540A1
; GENERAL INFORMATION:
; APPLICANT: ZHANG, ZHONGHUA
; APPLICANT: DUBOIS, RAYMOND
; TITLE OF INVENTION: IDENTIFICATION OF NSAID-REGULATED GENES
; FILE REFERENCE: VBLT:012US
; CURRENT APPLICATION NUMBER: US/10/206,901B
; PRIOR FILING DATE: 2002-11-19
; PRIOR APPLICATION NUMBER: 60/308,370
; PRIOR FILING DATE: 2001-07-27
; NUMBER OF SEQ ID NOS: 57
; SOFTWARE: Patent In Ver. 2.1
; SEQ ID NO 21
; LENGTH: 692
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-206-901B-21

Alignment Scores:
Pred. No.: 6.46e-30 Length: 692
Score: 374.50 Matches: 86
Percent Similarity: 62.44% Conservative: 42
Best Local Similarity: 41.95% Mismatches: 60
Query Match: 28.39% Indels: 17
DB: 15 Gaps: 7

US-10-088-830-2 (1-261) x US-10-206-901B-21 (1-692)

QY 50 LysAspArgLysLysGluLysAlaAlaProArgIleThrGlyTrpGlyLeuArgGlu 69
Db 55 CGCAACAGGAAAGGAGAGAAAG-----AATGCAAGGGCTACGGCAT 96
QY 70 TyrSerLysIleValCysGluLysValGluAlaLysGlyArgThrThrTyrAsnGluVal 89
Db 97 TTTCTCATCAAGTCTTGGAGAGGTGCGAGGAAGGAGGAGGACCCCTCTTACACAGAGTG 156
QY 90 AlaAspGluIleTyrSerGluLeuLysSerMet---AlaHisIle-----GlyGln 105
Db 157 GCAGACGAGCTGTTGGAGGTTCAGTGTGCGCAACACCATCTTACCAACAGGTCA 216
QY 106 GlyPheAspGluLysAsnIleArgArgValTyrAspAlaPheAsnValLeuIleAla 125
Db 217 GCTTATGACCAAAAAACATAGACGGCGGTACAGCGCTTAAACGTGCTAATGGCC 276
QY 126 LeuArgValIleAlaLysGluLysGluIleArgTyrMetGlyLeuSerAsnTyrArg 145
```


11115 Page Blank (uspto)